

UNIVERSITÀ DEGLI STUDI DI PADOVA
FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI
CORSO DI LAUREA SPECIALISTICA IN FISICA



Identification of gravitational wave signatures
with detector networks
by exploiting the response geometry
and wavelet packet decomposition

Relatore: Prof. Massimo Cerdonio Chiaromonte
Correlatore: Dott. Antonello Ortolan

Laureando: Tito Dal Canton
Matricola: 544356

ANNO ACCADEMICO 2007/2008

© 2008 Tito Dal Canton · <http://www.dalcanton.it/tito>

This work is licensed under the Creative Commons attribution–non commercial–share alike 3.0 license. To view a copy of this license, visit the web site

<http://creativecommons.org/licenses/by-nc-sa/3.0>

or send a letter to

Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.

Contents

Introduction	5
1 General relativity and gravitational waves	9
1.1 Newton's gravitation and general relativity	9
1.2 Einstein's field equations	11
1.3 Gravitational waves	12
1.3.1 Linearization of field equations	12
1.3.2 Relevant gw properties	13
1.4 Gw effect and detection	14
1.4.1 Resonant detectors	17
1.4.2 Interferometers	18
1.5 Gw generation and sources	19
1.5.1 Periodic sources	21
1.5.2 Bursts	22
1.5.3 Stochastic background	23
2 Experimental measurement of gws	25
2.1 Measurements in modern physics experiments	25
2.2 Gw detectors as linear systems	26
2.3 Sensitivity curve of a gw detector	27
2.3.1 Intrinsic noise from the fluctuation-dissipation theorem	28
2.3.2 Spurious signals from the external environment	30
2.4 Complete output of a gw detector	32
3 Data analysis in a detector network	33
3.1 Signal properties and data analysis requirements	33
3.2 Signal detection as hypothesis testing	35
3.3 Discrimination and multiple hypotheses test	39
3.4 Networks of gw detectors	40
3.4.1 Network response to gws	44
3.4.2 Inverse problem and network conditioning	48
3.4.3 Actual and ideal networks	52
3.4.4 Geometrical interpretation of the network response	56
3.5 Detection and discrimination through geometric projections	60
3.5.1 Statistics of E_P and E_Q	62
3.5.2 Time-frequency multiresolution analysis	67
3.5.3 The PQ plane	72

4	Implementation and Monte Carlo simulations	77
4.1	Prototype of the analysis algorithm	77
4.2	Simulation parameters	78
4.2.1	Detector networks and output signals	78
4.2.2	Discretization of the sky	79
4.2.3	Burst waveforms	80
4.3	Results of the Monte Carlo simulations	81
5	Conclusions and future perspectives	83
	Bibliography	85
	Acknowledgments	91

Introduction

Gravitational waves are the radiative component of the gravitational field, propagating at the same speed of light in vacuum and originated by the re-arrangements of energy/matter distributions [1, 2]. Their existence is prescribed not only by Einstein's general theory of relativity, but also by any theory of gravity based on a metric tensor. Although their existence was proven indirectly by observations of binary pulsar orbits (e.g. the PSR1913+16 [3] and the PSRJ0737-3039A/B [4]), they have never been directly detected in laboratories. Their direct detection will be important for different reasons.

Regarding the physics of gravity, testing the existence and the properties of gravitational radiation will provide another strong, independent confirmation of general relativity. It will also, in principle, provide tests for alternative theories of gravity, such as the Brans-Dicke theory predicting scalar gravitational waves [1].

The most exciting follow-up of gravitational wave detection, however, will be its application to future research in astrophysics and cosmology. In fact, compact astrophysical objects like neutron stars and black holes are expected to radiate large amounts of energy in gravitational waves [5, 6, 7, 8, 9, 10] which, if detected, can provide many details about the objects themselves. Due to the nature of their generation, gravitational signals would give complementary measurements with respect to electromagnetic observations, providing access to extreme phenomena taking place inside massive astrophysical objects, impossible to reproduce in laboratories and inaccessible to current instruments [11]. Furthermore, electromagnetic radiation is emitted by the collective, incoherent motion of individual atoms, while gravitational radiation arises from coherent, bulk motion of matter. Another advantage is the weak gravitational wave coupling with matter (the coupling constant is Newton's gravitational constant $G \simeq 7 \cdot 10^{-11} \text{ N m}^2 \text{ kg}^{-2}$). In fact, unlike light, gravitational radiation is only slightly attenuated by matter, making it possible to observe very distant objects or, equivalently, gravitational signals from the early universe, possibly from its inflationary phase [5, 12, 13].

The small coupling between gravitational waves and matter, however, also makes experimental measurements a very challenging task [14]. In fact, differently from what happens in other fields of experimental physics, the signal is expected to have the same order of magnitude of the detector intrinsic noise fluctuations and of other "spurious" signals of environmental origin, that enter the instrument through various couplings (e.g. mechanical

and electrical). This makes the recognition of gravitational signals with no known waveform with a single detector impossible *in principle*, and requires a *network* of $M > 1$ detectors and joint data analysis procedures.

Moreover, the emission of strong gravitational radiation is a process requiring extreme conditions, i.e. large masses moving at relativistic velocities and confined in regions comparable in size to the Schwarzschild radius. It is therefore impossible, from a practical point of view, to produce measurable gravitational waves on Earth [14]. On the one hand, this implies that we can only calibrate our detectors by means of forces equivalent to gravitational waves as prescribed by general relativity; on the other hand, detectable signals will only come from astrophysical and cosmological sources, the only phenomena meeting such extreme requirements. This introduces the further problem of having no control on the signal source (the signal parameters, like the time of arrival, are unknown) and of requiring accurate models for the source itself (without models, detection of the signal is much more difficult). Again, this leads to the requirement of a detector network and coherent data analysis techniques.

A number of gravitational wave detectors are currently online, aiming at detecting waves from astrophysical and cosmological sources. The main experiments are three kilometer-baseline interferometric detectors (LIGO [15] and VIRGO [16]) and three resonant-bar detectors (AURIGA [19], EXPLORER [20] and NAUTILUS [21]).

The current data analysis procedures for detecting gravitational waves involve different methods that depend on the gravitational source. To detect impulsive gravitational waves from compact object mergers, supernova core collapses and gamma-ray burst engines, the so-called *incoherent* methods are used, looking for time coincidence between event lists provided by different detectors. An important effort in this direction is done by the IGEC collaboration and uses data from bar detectors [22, 23], while data from interferometers are analyzed with coincidence searches like the WaveBurst algorithm [59]. For inspirals and continuous signals, *coherent* methods are preferred, since the strong phase coherence of the waveforms can be exploited [66]. Coherent searches look for cross-correlations between different detectors or search for the maximum likelihood over the signal parameter space. They are mostly used by the LIGO and VIRGO groups to analyze data from interferometers.

Despite the intensive efforts, all data analyses ran produced negative results: the *null hypothesis*, i.e. the statement “no gravitational signal is present in the data”, still has not been rejected (see for example [23, 42, 43, 44, 45]). Since the measured event rate with current detectors is too low to be useful for physics, research is now aiming at improving the sensitivity of interferometric detectors, and will eventually lead, in the next years, to a new network of more sensitive (*advanced*) detectors, greatly extending the observable distance and thus hopefully increasing the rate of gravitational events to reasonable values (a few events per year) [24, 25]. Research is also going on for resonant-mass detectors, whose natural evolution is towards spherical and DUAL acoustic detectors [26, 27, 28]. The future detector network will also be enhanced by a completely new and independent

space-borne instrument (LISA—Laser Interferometer Space Antenna) sensitive to a different class of gravitational signals impossible to detect on Earth, namely the 10^{-4} - 10^{-1} Hz low frequency band [29]. There are also research efforts aimed at devising new data analysis strategies, requiring fewer assumptions about sources and reasonable computational power. In particular, coherent methods look promising even for detecting impulsive gravitational waves, and they are being extended to this class of signals to overcome the limitations of incoherent searches [62, 63, 64, 60, 61].

This thesis belongs to the latter research field. Its purpose is the development and characterization of data analysis strategies for coherent detection of impulsive gravitational waves, and their discrimination from non-gravitational signals, with networks of detectors. In particular, we investigate possible ways to design a detection algorithm by exploiting the geometrical properties of the detectors response, rather than looking for coincidences between particular time-frequency structures in detector outputs or fitting the outputs to known signal templates.

The resulting detection algorithm is sensitive to signatures connected to the gravitational wave structure itself (in particular, its physical symmetries) and to the way detectors respond to the wave, without the need for specific assumptions about how the wave is actually generated. One expects robustness against unavoidable non-gravitational disturbances, possibly an improved detection efficiency with respect to existing methods, and reduced computational requirements.

We also review some problems connected to the existing network of detectors, we discuss how they limit the ability to gather information about gravitational signals and show how they could be reduced.

Monte Carlo numerical simulations are being performed, with support from the AURIGA group at the Legnaro National Laboratories (LNL) in Padova, to evaluate the performance of the proposed algorithm. Both existing and ideal detector networks are considered, to determine how a non-ideal network can spoil the performance of the method. For the first characterization, only simulated detector responses are being used.

The plan of the thesis is as follows.

In chapter 1 we give a brief introduction to Einstein's general theory of relativity, focusing on gravitational waves: their symmetry properties, the generation mechanism, their effects on matter and how these effects can be exploited for performing measurements. Expected astrophysical and cosmological sources are also briefly reviewed.

In chapter 2 we describe how the measurement process takes place in modern detectors, focusing on instrument response to gravitational waves and on the unavoidable contamination of the output by noise and spurious disturbances.

The data analysis devised to detect the presence of gravitational wave transients in the detector network outputs and to discriminate them from spurious signals is introduced in chapter 3. We give an introduction to hypothesis testing, summarize existing detection algorithms, describe the network response to gravitational waves and present the main results of the thesis.

In chapter 4 we provide the details of the algorithm implementation and explain the choice of detector networks and other simulation parameters. Results of the Monte Carlo simulations we are running to validate the algorithm are finally presented.

The conclusions are drawn in chapter 5, together with some topics that deserve further investigation.

Chapter 1

General relativity and gravitational waves

1.1 Newton's gravitation and general relativity

At the beginning of the 20th century, Newtonian dynamics and gravitation formed the best theoretic framework accounting for gravitational effects [32]. Their mathematical apparatus consists basically of the force-acceleration relation and the inverse-square law for the gravitational interaction

$$\mathbf{f} = m \frac{d^2 \mathbf{r}}{dt^2}, \quad \mathbf{f}_g(\mathbf{r}) = G \frac{mM}{r^2}, \quad (1.1)$$

where \mathbf{f} is the force, m is the mass, \mathbf{r} is the position vector measured in an inertial reference frame and \mathbf{f}_g is the gravitational force due to a second mass M . These reasonably simple equations allowed to predict the motion of the moon and planets in the solar system with unprecedented accuracy, leading also to the discovery of Neptune and other celestial bodies in the solar system.

However, Newton's theory is incomplete under several aspects [1, 32]. From an experimental point of view, for example, it fails to explain the precessing orbit of Mercury, requiring ad-hoc adjustments, such as the introduction of other celestial bodies, which are contradicted by observations. Regarding more theoretical aspects, the severe flaw of Newton's equations is that they only hold in a set of "privileged" frames of reference, i.e. the *inertial* ones. To support this idea and define this special class of frames, one is forced to postulate the existence of an "absolute" space. Moreover, Newton's theory makes no attempt to explain the physical origin of the *inertial forces* that an observer experiences when its reference frame is accelerated. Finally, from the introduction of Einstein's special theory of relativity in 1905, it became immediately clear that Newtonian gravitation, being based on an instantaneous action, was a non-relativistic theory and so it had to be revised.

Ten years after formulating special relativity, Einstein himself managed to complete this remaining task, making a significant breakthrough in our understanding of gravitation by means of the general theory of relativity,

or general relativity in short [33]. In his search for a better description of gravitation, Einstein was inspired by several principles, both qualitative and quantitative, of paramount importance and generality. The most important ones are now quickly reviewed [30].

The so-called *Mach's principle* states that inertial effects, including the existence of inertial reference frames, are not due to the presence of an absolute space, but are fixed by the mass distribution of neighboring objects. Therefore, the concept of inertial effect only makes sense with respect to neighboring objects. Then, in our universe one expects inertial reference frames to be at rest, or moving at constant relative velocity, with respect to the fixed stars, where most of the existing matter is thought to concentrate. This idea is supported by observations, as the fixed stars appear at rest (not rotating) with respect to inertial reference frames available on Earth. It is worth noting that Mach's point of view involves a coupling between all existing mass; this interaction is the origin of inertial forces.

The *equivalence principle* takes its origin from a number of observations, first done by Galileo Galilei, showing the numerical equivalence of inertial and gravitational mass. It is the quantitative statement that the motion of a particle experiencing only gravitational interactions does not depend on the particle composition, mass or other characteristics. That is, under gravitational interactions, all particles fall in the same way. This extremely important observation is now verified with very high accuracy [34] and is considered the key idea at the base of general relativity. An interesting reformulation of the principle is the statement that all gravitational effects are locally canceled whenever one chooses a free-fall reference frame, i.e. a frame at rest with particles falling in the gravitational field itself. Another form of the principle is that everything couples to the gravitational interaction: every form of energy, including mass, interacts in the same way with every other physical entity.

The *general covariance principle* is necessary whenever one wants to generalize the principle of special relativity to all observers, both inertial and non-inertial. Requiring the equivalence of all observers means that the equations of physics must hold in every coordinate system and frame of reference. This strong requirement implies that physical equations have to be expressed in tensorial form; this is the essence of general covariance. This principle was considered a fundamental assumption by Einstein, who stated that all observers, not only inertial ones, should be able to discover the laws of physics.

The well-known *correspondence principle* states that each new theory must be compatible with previous descriptions under opportune conditions: general relativity has to reduce to special relativity in absence of gravitational interactions, and to Newtonian gravitation in the limit of weak gravitational interactions and small velocities.

Note that the formulation of those principles differs from author to author; for example, some reduce them to two forms of the equivalence principle, the *strong equivalence principle* (general covariance) and the *weak equivalence principle* (equivalence between inertial and gravitational mass).

The crucial consequence of the principles is that, in general relativity,

spacetime must be identified and described with a curved Riemannian manifold [34]: gravitational effects are due to geometrical properties of spacetime, namely its curvature, represented by the Riemann curvature tensor $R_{\mu\nu\rho\sigma}$. Physical laws are represented by equations containing tensorial quantities, and involve tensorial algebra and calculus over the spacetime manifold. This accounts for general covariance. As in special relativity, motion of free particles and electromagnetic radiation is described by geodesic curves, whose trajectories are determined by the underlying spacetime geometry through the metric tensor $g_{\mu\nu}$. As geodesics are independent from the properties of particles and light rays, the equivalence principle is also satisfied. With respect to Newton's gravitation, the metric tensor $g_{\mu\nu}$ plays the role of the gravitational potential; therefore, the field equations involve the Riemann curvature tensor, a quantity which contains combinations of the second derivatives of $g_{\mu\nu}$. The field equations establish the spacetime geometry and its connection to the source of the gravitational field, which Einstein set equal to the stress-energy tensor $T_{\mu\nu}$ of special relativity. This seems to satisfy also Mach's principle, since inertial and gravitational effects (described by the spacetime geometry) are determined by the mass distribution (described by the stress-energy tensor).

Today, after a century since its formulation, general relativity is considered the most complete theory of gravitation at the macroscopic scale [32]. Thanks to astrophysical data collected in the last years, several predictions have been experimentally tested with good accuracy, even in the strong-field regime [3, 4, 34]. The constraints put by experiments on several parameters strongly favor general relativity over the other metric theories of gravitation. The theory has already influenced technology, as the widely used Global Positioning System needs to take into account general relativistic effects to operate correctly [35]. However, important predictions still lack a *direct* experimental confirmation, notably gravitational radiation [23, 42, 43, 44, 45]. Furthermore, there are also several open questions of theoretical nature, e.g. the relation between general relativity and Mach's principle [36].

1.2 Einstein's field equations

The complete Einstein's field equations and some of their properties are now briefly reviewed. The equations can be written as [1, 2]

$$R_{\mu\nu} - \frac{1}{2}R g_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu}, \quad (1.2)$$

where $R_{\mu\nu} = R^\rho{}_{\mu\rho\nu}$ is the Ricci tensor, $R = R^\mu{}_\mu$ is the scalar curvature, $g_{\mu\nu}$ is the spacetime metric tensor, $T_{\mu\nu}$ is the stress-energy tensor and $c \simeq 3 \cdot 10^8 \text{ m s}^{-1}$ is the speed of light in vacuum.

Physically, these equations couple the spacetime geometry to the energy/matter fields [37]. In fact, the left-hand side contains quantities describing the spacetime geometry, i.e. curvature and metric tensors, while the right-hand side contains all non-gravitational contributions like matter and electromagnetic fields. The coupling constant $8\pi G/c^4 \simeq 2 \cdot 10^{-43} \text{ N}$ fixes the

amount of energy density required to produce a unitary spacetime curvature. This amount is enormous and therefore under many conditions, including most of the solar system, spacetime can be considered flat to a good approximation. Note that, as the equations are between tensor quantities, they are invariant in form under coordinate transformations as required by general covariance. In general, these equations fix at the same time the evolution of the spacetime geometry due to the energy distribution and the motion of energy due to spacetime geometry, and this complicated interaction can not be decoupled. However, in presence of simplified configurations, the equations can be used to determine the spacetime metric given the source stress-energy tensor, i.e. the energy distribution.

From a mathematical point of view, the field equations are 10 second order, non-linear, coupled, partial differential equations involving 20 quantities (10 components of $g_{\mu\nu}$ and 10 of $T_{\mu\nu}$). The non-linearity and coupling of the equations makes the theory very difficult to handle. For example, the principle of superposition does not hold: linear combinations of solutions are not, in general, solutions themselves. Without this simplification it is very difficult to find exact solutions.

Nevertheless, there are very general and interesting classifications of the solutions [30]. An example is the Petrov classification, which groups the possible vacuum solutions (i.e. $T_{\mu\nu} = 0$) into a hierarchy of six “types”, depending on the algebraic properties of the Riemann tensor. The physical interpretation is that different types correspond to different classes of physical configurations: for example, one of the types groups spacetimes of isolated massive objects, while other types represent radiative solutions. Despite the complexity of the equations, for very symmetric configurations important analytical solutions have also been found and studied [30]. These include the Schwarzschild spacetime, i.e. the metric in presence of a non-rotating, electrically neutral mass distribution with spherical symmetry; the Kerr spacetime, i.e. the generalization of the Schwarzschild solution to a rotating mass distribution; the Reissner-Nordström spacetime, i.e. the further generalization to a charged and rotating spherical mass distribution. These solutions are very useful in modeling spherical distributions of mass, such as stars and black holes. When strong symmetries are not available, but the fields can be considered weak or the motion is non-relativistic, the equations are usually studied through series expansion, yielding the so-called *post-Newtonian* and *post-Minkowskian* approximations. A similar approximation leads to the linearized radiative solutions, i.e. gravitational waves, which will be introduced now.

1.3 Gravitational waves

1.3.1 Linearization of field equations

Consider an empty region of the spacetime far from large mass distributions, where the field equations reduce to $R_{\mu\nu} = 0$ and the geometry is described in first approximation by the flat Minkowski metric $\eta_{\mu\nu} = \text{diag}(1, -1, -1, -1)$. Assuming the presence of a small metric perturbation $h_{\mu\nu}$, so that $g_{\mu\nu} =$

$\eta_{\mu\nu} + h_{\mu\nu}$, we are interested in the evolution of $h_{\mu\nu}$. By plugging this perturbed metric into the field equations, neglecting terms of second and higher order in $h_{\mu\nu}$, and choosing the transverse and traceless gauge (*TT gauge*) defined by $h^\mu{}_\mu = \partial_\mu h^\mu{}_\nu = 0$, it can be shown that the equations reduce to $\square h_{\mu\nu} = 0$, where $\square = \eta^{\mu\nu} \partial_\mu \partial_\nu = \nabla^2 - c^{-2} \frac{\partial^2}{\partial t^2}$ is the d'Alembert operator [30]. This is the famous wave equation in vacuum and it is well known that, in general, its solutions represent waves propagating with finite velocity. In this context, the waves play the role of perturbations in the flat Minkowsky metric, and they are called *gravitational waves* (gws). It can be argued that a non-vanishing perturbation $h_{\mu\nu}$ is not necessarily indicative of a non-vanishing gravitational field, so the wave equation may not describe real gravitational effects. However, one can demonstrate that a similar wave equation holds also for the components of the Riemann curvature tensor. Therefore, gws are actually perturbations of the spacetime *curvature*, and consequently of the corresponding metric tensor. Furthermore, one can show that gws transport energy, linear and angular momentum, so they must represent a physical effect [31].

By choosing the TT gauge and orienting the coordinate system with the z axis parallel to the wave vector, it can be shown that only 2 of the 16 $h_{\mu\nu}$ components represent physically significant quantities. Moreover, in the same gauge and coordinate system, a generic plane gravitational wave can be described by the 3×3 tensor

$$\mathbf{h}(ct - z) = h_+(ct - z) \mathbf{e}_{TT}^+ + h_\times(ct - z) \mathbf{e}_{TT}^\times, \quad (1.3)$$

where h_+ and h_\times are two time-dependent amplitudes and

$$\mathbf{e}_{TT}^+ = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{e}_{TT}^\times = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad (1.4)$$

1.3.2 Relevant gw properties

The description of gws by means of the linearized general relativity involves a number of mathematical properties with interesting physical consequences [30], which will now be briefly reported.

First of all, from the wave equation it can be immediately recognized that the propagation speed of gws in vacuum is c , exactly like electromagnetic radiation. This is consistent with the principle of relativity: gravitational interactions are forced to propagate at a *finite speed*, like any other physical interaction.

As the perturbation is represented by a metric term, gws are described by a *symmetric tensor of order two*. This is equivalent to saying that they are a *spin-2 field*. This is a notable difference from the electromagnetic radiation, which is represented by a spin-1 field.

The general solution depicted above shows that a plane gw has only two “internal” degrees of freedom, represented by the two amplitudes $h_+(t)$ and $h_\times(t)$. These degrees of freedom are associated with two possible polarization states, commonly known as *plus* and *cross*, which correspond to the

tensors e^+ and e^\times once the TT coordinate system has been chosen. One can say that the gw spin-2 field possesses two physical helicity states, *left* and *right*, related to the plus and cross states through a simple base transformation [1]. These states differ by a 45° rotation about the wave vector, as it can be seen by applying such a rotation to the two tensors e^+ , e^\times . In principle, with proper amplitude and phase assigned to $h_+(t)$ and $h_\times(t)$, a gw can realize every possible linear combination of the polarization states, including linearly, circularly and elliptically polarized radiation, in full analogy with the electromagnetic field. Once the coordinate system is fixed, for polarized radiation it is useful to define a *polarization angle* $\psi = \tan^{-1}(h_\times/h_+)$, which will be constant for linear polarization and time-dependent for elliptical and circular polarization.

Another general property which can be evinced from the \mathbf{h} expression is $h_{ij}z^j = 0$, i.e. the wave tensor has a null eigenvalue in the direction of propagation. This holds in every coordinate system. This has the important physical implication that the perturbation is always orthogonal to the wave vector, i.e. it is *transverse*, another analogy with electromagnetic waves.

Finally, it can be readily seen that $\text{tr}(\mathbf{h}) = 0$, i.e. \mathbf{h} is a *traceless* tensor. As the trace is invariant under coordinate changes, this is true in any coordinate system. The physical meaning of this algebraic property is that a gw induces geometrical deformations that do not change volumes.

1.4 Gw effect and detection

A crucial assertion of general relativity is that the motion of a free-falling particle is represented by a geodesic curve of the spacetime manifold. Once the coordinate system is fixed and the geodesic curve is expressed in parametric form as $\xi^\mu(\tau)$, where τ is the geodesic proper time, it can be shown that [1]

$$\frac{d^2\xi^\mu}{d\tau^2} + \Gamma^\mu_{\nu\rho} \frac{d\xi^\nu}{d\tau} \frac{d\xi^\rho}{d\tau} = 0, \quad (1.5)$$

where $\Gamma^\mu_{\nu\rho}$ are the connection coefficients, quantities related to the metric and its first derivative. This differential equation is called *geodesic equation* and physically represents the particle equation of motion in the chosen coordinate system.

As is well known from the equivalence principle, gravitational effects can not be revealed in the geodesic of a single free-falling particle, but they can be detected by measuring deviations between the geodesics of neighboring particles. In fact, given two nearby geodesics $\xi^\mu(\tau)$ and $\xi^\mu(\tau) + \varepsilon^\mu(\tau)$, the equation of motion for the small separation vector ε^μ can be approximated, to first order in ε^μ , by the *geodesic deviation* equation [1]

$$\frac{D^2\varepsilon^\mu}{D\tau^2} = R^\mu_{\nu\lambda\rho} \varepsilon^\lambda \frac{d\xi^\nu}{d\tau} \frac{d\xi^\rho}{d\tau}, \quad (1.6)$$

where $\frac{D\varepsilon^\mu}{D\tau} = \frac{d\varepsilon^\mu}{d\tau} + \Gamma^\mu_{\nu\rho} \frac{d\xi^\nu}{d\tau} \varepsilon^\rho$ represents the covariant derivative along the curve $\xi^\mu(\tau)$. As the equation involves the Riemann curvature tensor, ε^μ is

clearly driven by gravitational effects and not by coordinate ones (e.g. inertial forces in accelerated systems). It is said that the two particles experience a *tidal deviation*. In the case of gws, “nearby geodesics” must be intended with respect to gw wavelength, so the geodesic deviation approximation holds whenever the physical system containing the free-falling test particles (i.e. the gw detector) is small with respect to the characteristic gw wavelength. It will be shown later that for existing gw detectors this is in many cases a very good approximation.

To intuitively discuss general relativity measurements, we need the reference frame where physics most resembles the classical Newtonian theory, the so-called *Fermi normal coordinates* (FNC). The FNC origin is attached to the center of mass of the two particles (or, in general, of the gw detector), its axes are orthogonal and aligned to three gyroscopes, and the time coordinate is given by a clock at rest in the origin [2]. In this system $g_{\mu\nu}$ is constant up to its first derivative and the space components of eq. 1.6 read

$$\frac{d^2 \varepsilon^i}{d\tau^2} = R^i{}_{0j0} \varepsilon^j = \frac{1}{2} \frac{d^2 h^i{}_j}{d\tau^2} \varepsilon^j. \quad (1.7)$$

Using expression 1.3, eq. 1.7 reduces to an ordinary Newtonian equation of motion

$$\ddot{\varepsilon}_i = \frac{1}{2} \ddot{h}_{ij}(t) \varepsilon_j = \frac{1}{2} \left[\ddot{h}_+(t) e_{ij}^+ + \ddot{h}_\times(t) e_{ij}^\times \right] \varepsilon_j. \quad (1.8)$$

Consequently, in the coordinate system of the laboratory, two nearby free-falling particles experience a relative tidal acceleration, with magnitude proportional to the second time derivative of the wave amplitude and direction determined by the polarization tensors e^+ and e^\times . The effect can be depicted by considering a ring of free particles (figure 1.1). A gw impinging along the ring axis causes a geometrical distortion of the ring shape, modulated by the gw amplitude. It is easy to show that the general gw properties are inherited by e^+ and e^\times and therefore manifest themselves in the motion of the ring; for example, the transverse character of the gw causes absence of motion in the direction of the wave vector, while the spin-2 character is responsible for the elliptical shape of the deformation and for the fact that the two polarizations differ by a 45° rotation about the wave vector.

Eq. 1.8 suggests that the gw amplitude can be inferred by measuring the distance between nearby point masses or, more concretely, the displacement of a volume element from its equilibrium in the laboratory frame [14]. However, as measurement instruments can only output a single number, what can be actually measured is the projection of the mass motion $\varepsilon(t)$ along the line joining the two particles. The gw measurement process is therefore represented by the saturation of the gw tensor \mathbf{h} with a constant tensor \mathbf{D} (*detector tensor*), determined by the rest position of the test masses with respect to the wave vector \mathbf{k} and thus containing all information about the detector geometry. The observable quantity $s_{gw}(t)$ is consequently a linear combination of the polarization amplitudes:

$$\begin{aligned} s_{gw}(t) &= h_{ij}(t) D^{ij} = \\ &= h_+(t) e_{ij}^+ D^{ij} + h_\times(t) e_{ij}^\times D^{ij} = \\ &= h_+(t) F^+(\mathbf{k}) + h_\times(t) F^\times(\mathbf{k}). \end{aligned} \quad (1.9)$$

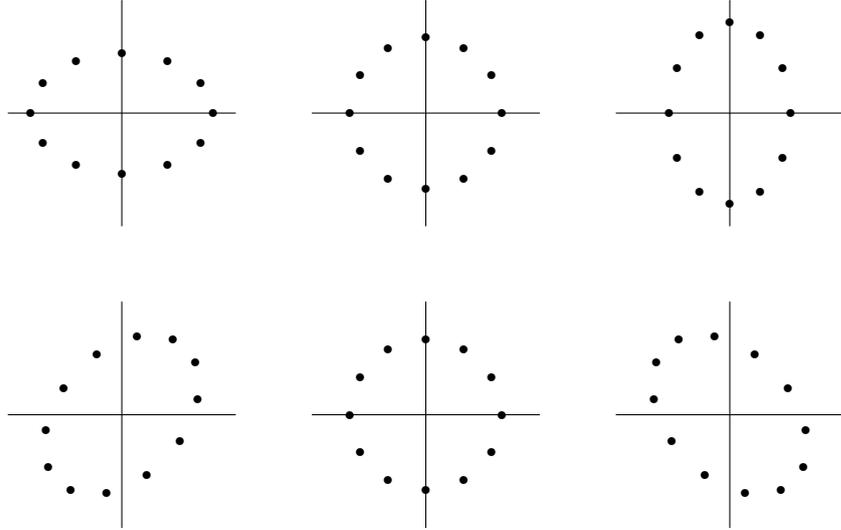


Figure 1.1: deformation of a ring of free particles induced by a gw, with plus polarization (upper row) and cross polarization (lower row). The left column shows a positive wave amplitude ($h > 0$), the middle column a null amplitude (no deformation) and the right column a negative amplitude ($h < 0$).

The coefficients

$$\begin{aligned} F^+(\mathbf{k}) &= e_{ij}^+ D^{ij} = \text{tr}(\mathbf{e}^+ \mathbf{D}) \\ F^\times(\mathbf{k}) &= e_{ij}^\times D^{ij} = \text{tr}(\mathbf{e}^\times \mathbf{D}) \end{aligned} \quad (1.10)$$

form the detector *antenna pattern* and represent geometric factors fixed by the relative orientation between the detector and the wave vector \mathbf{k} . Clearly, the polarization matrices have to be expressed in the detector coordinate system, that is

$$\begin{aligned} \mathbf{e}^+ &= \mathbf{E} \mathbf{e}_{TT}^+ \mathbf{E}^T \\ \mathbf{e}^\times &= \mathbf{E} \mathbf{e}_{TT}^\times \mathbf{E}^T, \end{aligned} \quad (1.11)$$

where \mathbf{e}_{TT}^+ and \mathbf{e}_{TT}^\times are those defined in the TT frame (eq. 1.4) and \mathbf{E} is a suitable \mathbf{k} -dependent Euler rotation matrix transporting the TT frame to the detector one [38]. F^+ and F^\times are usually normalized to be dimensionless numbers between -1 and $+1$, so they express the detector angular sensitivity to gws. For this reason they are often written as functions of the spherical-polar angles (ϑ, φ) describing the orientation of the gw source, so that the wave vector is given by

$$\mathbf{k} = -k \begin{bmatrix} \cos(\varphi) \sin(\vartheta) \\ \sin(\varphi) \sin(\vartheta) \\ \cos(\vartheta) \end{bmatrix} \quad (1.12)$$

(this notation will also be used in the rest of the thesis). In general F^+ , F^\times or both may vanish for some direction, so the detector may be insensitive

to gws for some combinations of direction and polarization. A polarization-independent angular sensitivity can be defined by averaging the antenna pattern function $F^+(\vartheta, \varphi)$ over all polarization angles, yielding

$$\widehat{F}(\vartheta, \varphi) = \sqrt{\frac{[F^+(\vartheta, \varphi)]^2 + [F^\times(\vartheta, \varphi)]^2}{2}}. \quad (1.13)$$

Zeros of this function represent directions for which the detector is completely blind to gws. It is worth noticing that the antenna pattern is remarkably determined by the form of the polarization states and so reflects the gw properties and symmetries. Other kind of waves, e.g. scalar gws predicted by extensions of general relativity, would be sought by the same detector with a different angular sensitivity [39].

Existing gw detectors implement eq. 1.8 with two main approaches, reviewed in the following subsections.

1.4.1 Resonant detectors

An approach to measure the tidal deviation between test particles consists in monitoring the deformation of a classical mechanical harmonic oscillator, i.e. two test masses joined by a restoring force whose magnitude is linear in the displacement from the rest position. The first attempt in this direction was made by Weber [40] and is still exploited by several detectors (AURIGA [19], NAUTILUS [21] and EXPLORER [20]). By studying the equation of motion of the system subject to the gw force, it can be shown that such detectors give the maximum response to gws in a rather narrow frequency band near the mechanical resonance.

In practice, setting up two test masses linked by an harmonic force is very challenging; fortunately, mechanical oscillation modes of elastic bodies are equivalent to harmonic oscillators in the neighborhood of their characteristic frequencies. Therefore, real resonant gw detectors monitor the energy of one (or several) oscillation modes of a large elastic body by means of suitable transducers that convert mechanical vibrations into electrical signals (figure 1.2). Historically, the elastic body is a cylinder-shaped bar and the chosen oscillation mode is the fundamental longitudinal one, a setup that resembles closely the theoretical two-mass oscillator [14]. Operating bars have a mass

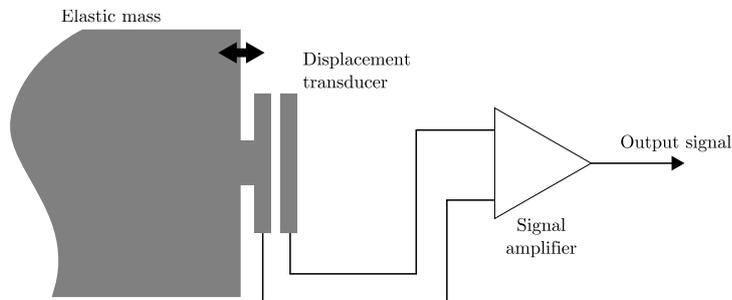


Figure 1.2: detecting gws with a resonant detector.

of $\sim 3 \cdot 10^3$ kg, length of ~ 2 m and fundamental mode at ~ 1 kHz. They thus exhibit high sensitivities in narrow bands around that frequency.

The geometry of a resonant-bar detector is described by a versor \mathbf{n} , parallel to the bar length. As the bar is equivalent to two test particles, the geodesic deviation equation leads to the detector tensor $D_{ij} = n_i n_j$. In a Cartesian coordinate system with the x axis aligned along \mathbf{n} , the antenna pattern functions then read

$$\begin{aligned} F^+(\vartheta, \varphi) &= \cos^2(\vartheta) \cos^2(\varphi) - \sin^2(\varphi) \\ F^\times(\vartheta, \varphi) &= \cos(\vartheta) \sin(2\varphi). \end{aligned} \tag{1.14}$$

They are plotted in figure 1.4. As evident from the \widehat{F} plot, they both vanish in the two opposite directions ($\vartheta = \pi/2, \varphi = 0$) and ($\vartheta = \pi/2, \varphi = \pi$), corresponding to the x axis, which we set parallel to the bar. A bar detector is therefore completely blind to waves travelling parallel to its length. On the contrary, the maximum sensitivity is to waves impinging perpendicularly to it.

1.4.2 Interferometers

The physical principle exploited by interferometric gw detectors is the interference between coherent light beams. The approach was first demonstrated by Forward [41]. The basic topology of interferometric detectors is depicted in figure 1.3. A coherent light beam is divided by a splitter in two beams, which travel to two far away, free-falling test masses. After being reflected, the beams rejoin in the beam splitter, forming an interference pattern in the optical detector (photodiode). The detector response can be used to reconstruct even tiny relative displacement of the test masses. In fact, as shown in figure 1.1, gws tend to stretch one of the arms while compressing the other, leading to a difference in the two optical paths and thus in the relative phases of the beams joining back in the beam splitter.

Practically, the free-fall condition of the test masses, beam splitter and light beams can be achieved approximately and only in a limited frequency band. This is accomplished using very complex suspension systems that decouple the mass motion from the external environment as much as possible, while ultra-high vacuum along all the optical paths limits light-gas interactions that would affect the phase measurement. Despite the technological difficulties of this approach, current interferometers (LIGO [15], VIRGO [16], GEO600 [17] and TAMA300 [18]) are orders of magnitude more sensitive than the original prototype and are both the most sensitive and most wideband available gw detectors [14], as will be discussed in chapter 2.

The geometry of an interferometer is described by two versors \mathbf{n} and \mathbf{m} , aligned along the arms. As, from a geometric point of view, each arm performs distance measurements between two test masses, the resulting detector tensor is $D_{ij} = \frac{1}{2}(n_i n_j - m_i m_j)$, resembling the combination of two resonant-bar detectors aligned along the arms. Although the arms need not be orthogonal, in operating detectors this is approximately the case, because this setup matches the quadrupolar symmetry of gws, improving the instrument sensitivity. In Cartesian coordinates with x and y along the arms,

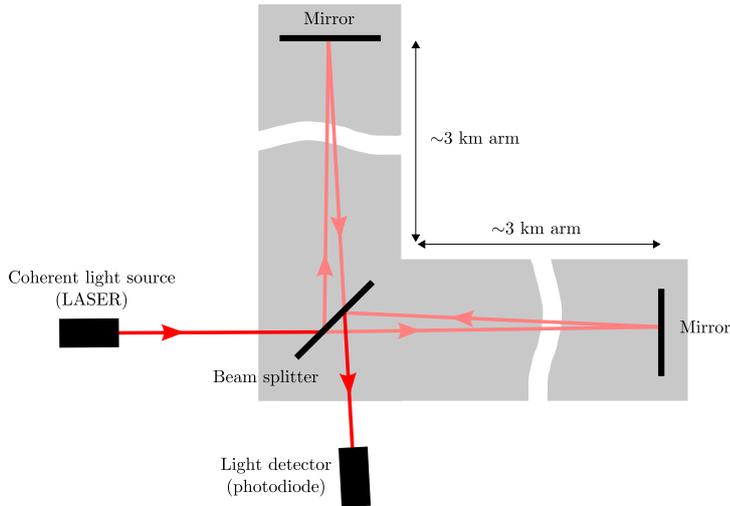


Figure 1.3: basic layout of an interferometric gw detector.

orthogonal-arm interferometers have the antenna pattern

$$\begin{aligned}
 F^+(\vartheta, \varphi) &= \frac{1}{2}(1 + \cos^2(\vartheta)) \cos(2\varphi) \\
 F^\times(\vartheta, \varphi) &= \cos(\vartheta) \sin(2\varphi),
 \end{aligned}
 \tag{1.15}$$

which is plotted in figure 1.4. Both functions vanish in the four directions given by $\vartheta = \pi/2$ and $\varphi = \{\frac{1}{4}\pi, \frac{3}{4}\pi, \frac{5}{4}\pi, \frac{7}{4}\pi\}$, the bisectors of the xy plane. Interferometers are therefore completely blind in four directions on the plane of the arms. On the other hand, they are most sensitive to waves coming perpendicularly to this plane, the condition of perfect match between the wave and instrument symmetries.

It can be easily seen that $\text{tr}(\mathbf{D}) = \|\mathbf{n}\|^2 - \|\mathbf{m}\|^2 = 0$: the detector tensor of an interferometer has a vanishing trace. This does not happen for resonant-bar detectors. This is yet another indication of the interferometer matching to the gw geometry, but it effectively prevents testing the vanishing trace of gws, as any non-null trace contribution is projected away from the instrument response. Indeed, this can be intuitively understood by remembering that the physical role of the trace is a “breathing” deformation (volume change). Such deformations can not evidently be sensed by an interferometer, as they represent common-mode length changes of the two arms.

1.5 Gw generation and sources

Formally, the mechanism giving rise to linearized gravitational radiation from a source is very similar to the electromagnetic case. In fact, one can prove that the right-hand side of the wave equation $\square h_{\mu\nu} = 0$, which vanishes in vacuum, represents the source term and contains the source stress-energy tensor. The solution is therefore the well-known *retarded potential*

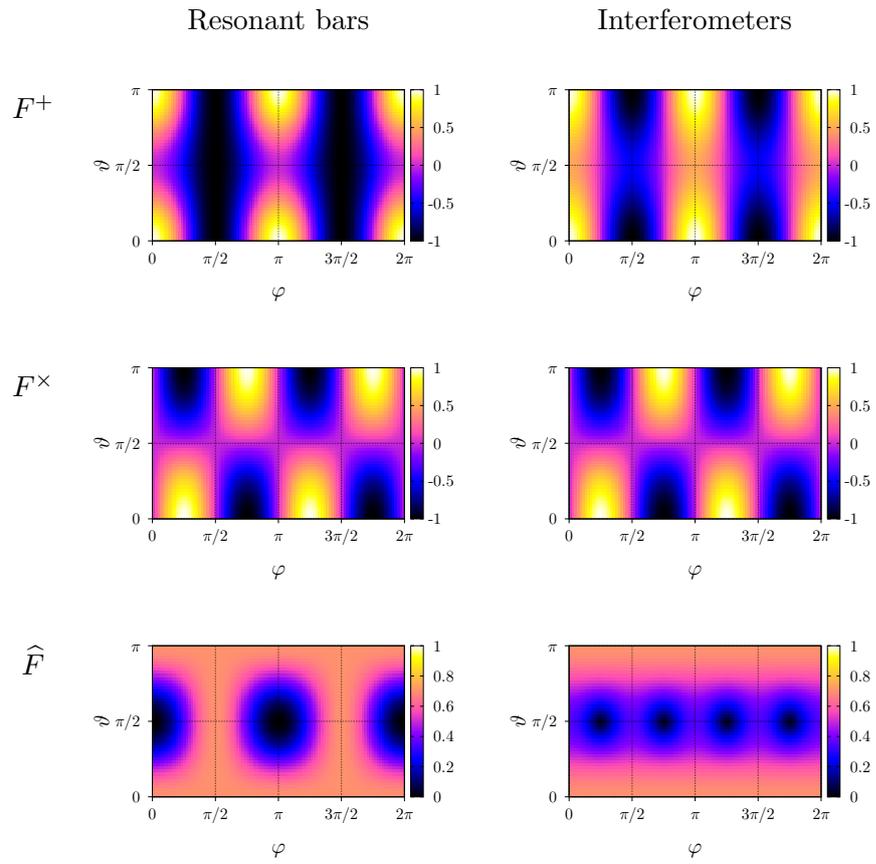


Figure 1.4: antenna pattern functions for the discussed detector schemes.

[1]

$$h_{\mu\nu}(\mathbf{x}, t) \propto \int \frac{S_{\mu\nu}(\mathbf{y}, t - |\mathbf{x} - \mathbf{y}|)}{|\mathbf{x} - \mathbf{y}|} d^3\mathbf{y}, \quad (1.16)$$

where $S_{\mu\nu} = T_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}T^\rho{}_\rho$. An important general result known as the *peeling theorem* [30] states that, moving away from an isolated source towards infinity, the various spacetime components of the Petrov hierarchy progressively “peel off”, leaving only the radiative solution at infinity in full analogy with electromagnetism, in which the space around the source divides into the “near” and “radiation” zones.

Gws propagating in the radiation zone can be expressed by expanding the source in multipole moments, which reveals another important distinction between gws and electromagnetic waves: the first multipole moment contributing to gravitational radiation is the *quadrupole*, associated with $l = 2$ spherical harmonics. This fact holds regardless of the source properties. Its physical reason is that the first two multipole moments represent the source mass and its center-of-mass linear momentum; these are strictly conserved for isolated sources, so their time dependence vanishes and they can not contribute to gravitational radiation [2]. Moreover, it can be shown that higher multipole moments are scaled by increasingly negative powers of c , so the dominant contribution is always by far the quadrupole.

Whenever the source also involves slow internal motion, the emitted radiation in the wave zone can be further approximated by the so-called *quadrupole formula* [32]

$$h_{ij}(\mathbf{x}, t) = \frac{G}{c^4} \frac{2}{r} \left(\frac{d^2 Q_{ij}}{dt^2} \right)_{t-r/c}^{TT}, \quad (1.17)$$

where $Q_{ij}(t) = \int \varrho(\mathbf{x}, t) (x_i x_j - \frac{1}{3} x^2 \delta_{ij}) d^3\mathbf{x}$ is the traceless quadrupole moment of the source mass density ϱ and the superscript TT means that only the transverse and traceless component of the quadrupole moment contributes to the radiation. This approximation is very good in the case of non-relativistic sources, but gives correct orders of magnitude even for relativistic ones. From the quadrupole formula one can approximate the total power radiated by the source (luminosity) as

$$\mathcal{L} = \frac{G}{5c^5} \frac{d^3 Q_{ij}}{dt^3} \frac{d^3 Q^{ij}}{dt^3}. \quad (1.18)$$

The factor $G/c^5 \simeq 3 \cdot 10^{-53} \text{ W}^{-1}$ strongly suppresses the emission: it is practically impossible to produce relevant gws on Earth. The only hope for significant time-variations of $\mathbf{Q}(t)$ is given by cosmic phenomena [14]. In literature, cosmic gw sources are usually classified into three categories, according to the very different time-frequency characteristics of the signals they emit [32]. The study of cosmic gw radiation is still an open problem, and the current status is reviewed in the following subsections.

1.5.1 Periodic sources

By definition, periodic (and quasi-periodic) gw signals are composed of a discrete set of frequencies whose evolution is negligible with respect to the

observation time. Clearly, these signals can be described with a limited number of parameters. For example, simple monochromatic signals can be modelled conveniently as

$$\begin{aligned} h_+(t) &= h_0^+ \cos(\omega t + \varphi_0^+), \\ h_\times(t) &= h_0^\times \cos(\omega t + \varphi_0^\times), \end{aligned} \tag{1.19}$$

where the amplitudes h_0^+ and h_0^\times , the angular frequency ω and the initial phases φ_0^+ , φ_0^\times are in general slowly varying functions of time.

Among the expected periodic sources there are quasi-stationary astrophysical systems [5] like binary stars [3, 4] and spinning compact objects exhibiting deviations from axial symmetry (e.g. pulsars and neutron stars) [6]. These sources can often be modelled by treating relativistic effects as perturbations of Newtonian motions: on the one hand, this allows to accurately predict the waveform of the emitted gw signal to arbitrary precision (e.g. with the quadrupole formula); on the other hand, the signal parameters (e.g. frequency) can be related to physical parameters of the radiating system, like mass and orbital period. Such systems, especially those in highly relativistic regimes, may also involve more complex behaviours like spin-orbit coupling and orbital precession, which influence the emitted gw signal. It is often said that the gw signal “maps” the spacetime metric in the neighborhood of the emitting system, extracting detailed information about its dynamics.

Though still uncertain, the expected number of quasi-periodic sources within the observable universe is very large. Most of the expected frequencies range approximately between 10^{-4} Hz and 10^3 Hz. Consequently, the wavelength spectrum extends above $\lambda_{min} = c/f_{max} \simeq 300$ km and the geodesic deviation used in existing ground-based detectors is a good approximation.

1.5.2 Bursts

A gw signal is classified as a *burst* if its amplitudes $h_+(t)$ and $h_\times(t)$ behave like transients (i.e. they are impulsive signals) and feature a time scale shorter than the observation time. Due to this broad definition, bursts can not be parametrized using a small number of parameters and must be described instead by specifying the full waveforms for $h_+(t)$ and $h_\times(t)$. In some cases, however, such precision is not necessary and bursts can be given rough descriptions, e.g. their amplitude and supporting region of the time-frequency plane.

Strong bursts are thought to originate from violent astrophysical processes, lasting for short times and involving large amounts of energy. This definition embraces an extreme variability of phenomena [5], including supernova and hypernova collapses [7], matter falling into black holes and coalescence of compact objects, in particular the final merging of binary system composed of neutron stars and black holes [8, 9, 10]. Presumably, there are coincidences between gw and gamma-ray bursts. Due to their highly relativistic nature, these phenomena involve general relativity in its most non-linear and strong regime: they can not be treated with perturbative methods. Furthermore, collapses and explosions are complex processes,

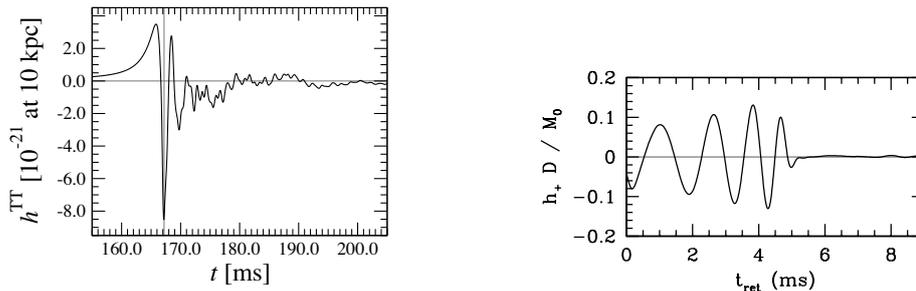


Figure 1.5: example gw bursts from numerical simulations. Left: supernova core collapse [46]. Right: final merger of a black hole/neutron star binary system [9].

highly dependent on the state equation of the participating matter which is still poorly known. Therefore it is impossible to analytically predict the emitted waveforms, but it is also hard to accurately estimate their rough parameters. Nevertheless, numerical relativity is now able to perform accurate simulations even in the strong-field regime [46, 10] and so represents a unique tool to study the resulting waveforms (figure 1.5).

Despite all the recent accomplishments, however, even the most elementary quantities like the expected burst event rates and wave magnitudes still rely on unsettled astrophysical observations and models: this leads to severe uncertainties, and the estimates still span several orders of magnitude: astrophysical bursts are expected to last from a few ms to several seconds, while their frequency content is thought to cover the kHz band. Therefore, ground-based detectors can safely apply the geodesic deviation approximation to detect bursts.

1.5.3 Stochastic background

Similarly to what happens for the cosmic microwave background, the existence of a non-deterministic (stochastic) gw signal filling the universe is expected [5]. As stochastic signals are characterized by specifying their statistics and spectral power density, a good description of this gw “background” is given by its relative, dimensionless spectral energy density

$$\Omega_{gw}(f) = \frac{1}{\varrho_c} \frac{d\varrho_{gw}}{d \ln(f)} \quad (1.20)$$

where ϱ_c is the critical density of the universe. This definition contains all information about the gw background when complemented by a few assumptions, i.e. isotropy, stationarity and Gaussianity of the fluctuations. It must be noted, however, that such reasonable requirements may not actually hold.

The presence of such a background is prescribed both by astrophysical considerations and cosmological models. From an astrophysical point of view, a contribution is expected from the superposition of many random, independent and unresolved gw sources of the kind depicted in the previous subsections: for example, an expected strong contribution is from galactic

and extra-galactic white-dwarf binaries. Several cosmological models also predict a “relic” stochastic gw background, though they are still speculative and affected by severe uncertainties [12]. The inflationary scenario predicts stochastic contributions originated as quantum fluctuations during the inflationary period. Other models predict a phase transition in the cooling early universe, producing expanding bubbles whose collisions could have radiated a great amount of energy as stochastic gws. Another contribution could be due to relativistic vibrations of primordial topological defects known as “cosmic strings” [13]. Supersymmetric theories predict the primordial excitation of scalar fields, also leading to stochastic gw radiation. Finally, some models regard the universe as a $(3 + 1)$ -dimensional “brane” embedded in a higher-dimensional space, and it could have been coherently excited to vibrate and give rise to a stochastic gw contribution.

Characteristics of the resulting astrophysical-cosmological stochastic gw signal are highly uncertain, but the covered bandwidth is expected to span several orders of magnitude (the geodesic deviation approximation can be used to detect only the long wavelength part of the spectrum). Electromagnetic observations, like the CMB anisotropies and millisecond pulsar timing, put severe constraints on the stochastic background intensity, at least in some frequency bands (e.g. $\Omega_{gw} < 10^{-8}$ at 10^{-8} Hz). However, there is still much to learn about the actual gw background and the final word can only be set by accurate observation and data analysis from future, more sensitive gw detectors.

Chapter 2

Experimental measurement of gws

2.1 Measurements in modern physics experiments

In modern instruments devoted to experimental physics, it has become common practice to represent the measured observable quantities in the digital domain, that is, as a finite set of finite-precision numbers. This choice has a number of advantages. For example, the digital domain is the same representation used by modern computers, which are invaluable tools for storing and analyzing the large amounts of data produced by the instruments. Once the measured data have been stored, filtering and analyzing it can be done straightforwardly through sophisticated digital signal processing techniques. These algorithms can be applied to the original data even multiple times and with different setups, allowing deep and differentiated analyses which can even run concurrently. In fact, the digital domain allows to store and retrieve information with minimum or no alteration at all, provided that the initial analog to digital conversion is performed with sufficient accuracy.

Modern gw detectors make no difference [14]. The mechanical or optical analog and time-dependent output of a gw detector is usually converted to a voltage or current by an opportune transducer, then amplified and finally fed to an analog-to-digital converter (ADC), which samples it at discrete time steps and outputs a finite-precision number $x[j]$ for each step, where j is an integer. The time series is possibly pre-processed and finally stored for later data analysis procedures. To result in minimum alterations in the information carried by the analog signal, the ADC should produce a number as linear as possible with respect to the input signal. Moreover, the output number must have a sufficient precision (i.e. number of bits) to introduce negligible quantization noise. Finally, the ADC sampling frequency must be high enough to include the full detector bandwidth, i.e. it must be at least twice as large, as required by the sampling theorem.

2.2 Gw detectors as linear systems

As discussed in chapter 1, gws are tensorial signals. However, due to the detector design, the measurement can only output a scalar quantity evolving in time (e.g. voltage) which is finally converted to digital samples. Therefore, the measurement process involves an unavoidable *projection* of the gw tensor $\mathbf{h}(t)$ against a certain detector response tensor \mathbf{D} , so that the experimental observable is actually $s_{gw}(t) = D_{ij} h^{ij}(t)$. As depicted in chapter 1, this is a linear combination of the input gw polarization amplitudes, with coefficients (antenna pattern) fixed by the relative orientation of the wave vector and detector geometry: the instrument responds linearly to the time dependency of the gw amplitude.

However, it is unphysical to assume that the scalar projection $s_{gw}(t)$ depicted above contributes directly to the signal available at the detector output port. In fact, physical detectors always exhibit a limited bandwidth: very low and very high frequency components of input signals are inaccessible to the detector output. Assuming the detector behaves as a *linear* and *time-invariant* system, such an effect can be modeled with an opportune band-passing frequency transfer function $G(f)$ applied to the scalar projection $s_{gw}(t)$. In the time domain, the contribution of the input gw signal to the detector output can therefore be written in terms of the convolution

$$s_{gw}(t) = \int_{-\infty}^{+\infty} g(t - \tau) s_{gw}(\tau) d\tau = \int_{-\infty}^{+\infty} g(t - \tau) D_{ij} h^{ij}(\tau) d\tau, \quad (2.1)$$

where $g(t)$ is the time-domain filter associated with the detector transfer function $G(f)$ through an inverse Fourier transform. Usually, $G(f)$ also contains the scale factor converting the dimensionless spacetime strain to the units of the output signal¹. Though this “complete” transfer function can usually be estimated by lumped element models or with the aid of numerical simulations as in [47], in practice it must be evaluated experimentally by means of some (continuous) *calibration* procedure, which is unavoidably affected by errors (relative calibration errors in $|G(f)|$ are estimated around 5–10% for current interferometers [48, 49]). Note that an ideal experimental calibration of the transfer function should take place by feeding a known gw input signal to the detector and measuring the output, as is usually done in other fields of experimental physics. In the case of gw signals this is clearly unfeasible because, as seen in chapter 1, we are unable to produce sufficiently strong signals on Earth. Experimenters must therefore resort to calibrated forces that mimic the gw effect expected from general relativity [48].

The first data analysis step performed on the measured output $x(t)$ should be the *deconvolution* with the detector transfer function $G(f)$, to reconstruct the dimensionless spacetime strain projection $s_{gw}(t)$ at the detector input. This is the reason why $G(f)$ must be estimated accurately.

¹We will adopt the convention of writing a generic detector-generated signal as ‘ $a(t)$ ’ when its units are those of the detector output and as ‘ $a(t)$ ’ when it is expressed as a gw-like dimensionless signal at the detector input. Therefore $a(t) = \int_{-\infty}^{+\infty} g(t - \tau) a(\tau) d\tau$.

As the convolution plays the role of a projection, it is worth noting that the detector acts on the input gw as a pair of independent projections: the first converts the tensorial quantity to a scalar one, while the second selects the finite observable band in the time-frequency plane. How these projections operate depends on the detector characteristics.

Since in general projections involve an information loss, recovering all the information about an unknown gw signal from the output of a single detector is *a priori* impossible, i.e. a single detector can not perform the full signal deconvolution. It will be clear from chapter 3 how a natural solution to this problem is the exploitation of a detector network.

2.3 Sensitivity curve of a gw detector

The output of every experimental apparatus designed to study a certain physical phenomenon is contaminated by fluctuations originated by other physical processes. In some fields of experimental physics, the fluctuations can be made weak enough to inflict an effectively negligible error to the measured quantity. In the case of gw experiments, signals of non-gravitational origin continuously contribute to the detector output. However, despite the great technological achievements of the last years, for current detectors their magnitudes are still comparable to or greater than the expected gw signals, so that they effectively *dominate* the detector output $x(t)$ [14]. Therefore, their characterization is of paramount importance and their presence has to be taken into account from the design of the instrument to the final data analysis procedures.

In general the non-gravitational contribution is stochastic and it can be described by the statistics $P(x)$ and by the frequency power spectrum $S_{xx}(f)$ of its fluctuations. However, it is unsuitable to give this description directly in units of the detector output. Instead, once the deconvolution with the transfer function $G(f)$ has been performed on $x(t)$, and the dimensionless gw strain has been reconstructed, the non-gravitational contribution is expressed in terms of an equivalent gw signal at the detector input. The frequency power spectrum of the equivalent “gw noise”, defined as

$$S_{hh}(f) = \frac{S_{xx}(f)}{|G(f)|^2}, \quad (2.2)$$

forms the so-called *sensitivity curve* of the detector (sensitivity in short), a useful tool to characterize the instrument performance in the frequency domain. In fact, being expressed in units of dimensionless spacetime strain (per unit frequency band), it can be directly compared to hypothetical gw signals, thus providing a first-glance estimate of the signal-to-noise ratio per unit frequency band, namely $\varrho^2(f) = |H(f)|^2/S_{hh}(f)$ where $H(f)$ is the Fourier transform of the gw signal $h(t)$. Furthermore, the performances of different detectors can be readily compared just by looking at their sensitivities: the lower sensitivity curve determines the more sensitive instrument in the frequency band under consideration. Quantitatively, $S_{hh}(f)$ allows to directly evaluate the variance of the fluctuations σ^2 in a frequency band

$[f_0, f_1]$ by means of the integral $\sigma^2 = \int_{f_0}^{f_1} S_{hh}(f) df$. It is worth noting that the bandwidth of any gw detector is necessarily limited, as the detector noise overcomes any gw signal for both small and large values of f . On the other hand, $S_{hh}(f)$ reaches a minimum in the frequency band where the instrument has its highest sensitivity. This is evident from figure 2.1, showing the theoretical sensitivity curves of some existing and future detectors, calculated by superposition of the modelled noise contributions.

In general, fluctuations affecting the detector performances can be grouped into two classes by considering their physical origin: spontaneous (e.g. thermodynamical) fluctuations and signals due to the interaction of the detector with its physical environment. As the following subsections will show, this separation is justified by important differences between the two families.

2.3.1 Intrinsic noise from the fluctuation-dissipation theorem

According to thermodynamics, physical observables of any macroscopic system at finite temperature $T > 0$ are affected by spontaneous, stochastic fluctuations around their mean values [50]. Of course, a measurement instrument is a physical system at finite temperature on its own, and its output $x(t)$ clearly represents a physical observable. Therefore, $x(t)$ is affected by stochastic fluctuations $n(t)$, superimposed to the contribution we want to measure and originated from the instrument itself. This contribution is generally called the *intrinsic noise* of the instrument. In current gw detectors, gw signals are expected to be of the same order of magnitude of intrinsic fluctuations [14]. The accurate characterization of such fluctuations, in particular the study of their power spectrum $S_{nn}(f)$ and statistical distribution $P(n)$, is therefore of paramount importance.

For an isolated system at thermodynamical equilibrium, a very general result known as the *fluctuation-dissipation theorem* [50] states that $S_{nn}(f)$ is determined by the system generalized susceptibility to external stimuli $G(f)$ through

$$S_{nn}(f) \propto \frac{kT}{f} \Im[G(f)], \quad (2.3)$$

where $k \simeq 1 \cdot 10^{-23} \text{ J K}^{-1}$ is the Boltzmann's constant (actually, eq. 2.3 is the classical limit $kT \gg hf$ of a more general relation, valid also in the quantum regime). Applying eq. 2.3 to the gw case, $G(f)$ represents nothing more than the detector transfer function, which transfers the effect under measurement (gws) to the output observable $x(t)$. As the imaginary part of $G(f)$ describes how the thermodynamical system dissipates the energy introduced by the external force, fluctuations are therefore essentially determined by the detector dissipation mechanism. Moreover, the stochastic process $n(t)$ has zero mean by definition and it is *stationary*, i.e. it retains the same statistical properties over time. An evident consequence of eq. 2.3 is that the intrinsic noise reduction requires lowering the detector temperature and limiting its dissipations as much as possible. The fluctuation-dissipation theorem, and in particular eq. 2.3, play an important role in calculating the theoretical sensitivity curves shown in figure 2.1. The difference between

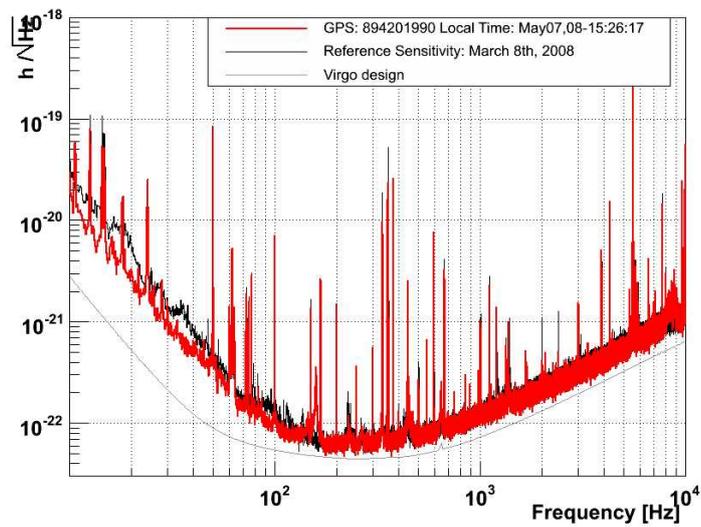
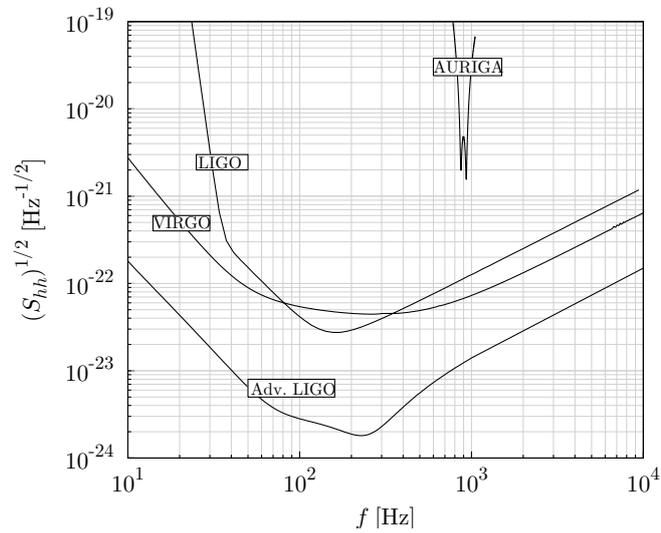


Figure 2.1: on the top, theoretical sensitivity curves for some current and future (advanced) detectors, showing the total noise expected from modelled sources [15, 16]. On the bottom, real sensitivity curve for the VIRGO interferometer [16]. Deviations from predictions are evident.

AURIGA and the other curves reflects the deep differences in the transfer functions of resonant bars and interferometers, notably the much larger bandwidth available with interferometers.

For isolated systems at equilibrium, the probability density $P(n)$, regarded as the distribution of $n(t)$ over the ensemble of systems at a fixed time instant, can also be predicted [50]. In fact, if x is a classical physical observable (i.e. quantum effects can be neglected) then its fluctuations about its mean value obey $P(n) = a e^{S(n)}$, where a is a normalization constant and S is the entropy of the system regarded as a function of n . The entropy is clearly maximum when n equals its mean value 0. Then, if the fluctuations are small, we can expand $S(n) \simeq S(0) - b n^2/2$, yielding

$$P(n) = a' e^{-b n^2/2} \quad (2.4)$$

which means that small fluctuations follow the *Gaussian* distribution when the detector is in thermal equilibrium.

The approximations that x behaves classically and that its fluctuations are small do generally hold quite well for gw detectors. However, whether such detectors can be considered isolated and in thermodynamical equilibrium is a much more delicate issue. The fluctuation-dissipation theorem and the Gaussian distribution may not hold strictly, i.e. they may apply only in limited frequency bands and for limited time intervals. Consequently, the validity of their predictions has to be verified continuously during the operation of the detectors. In practice this is done by comparing the estimated sensitivity curves with the theoretical models and monitoring the noise probability distribution. What is found is that the predictions do actually hold in limited frequency bands and time intervals, and that the intrinsic noise is a *quasi-stationary* process, i.e. its statistical parameters vary with time on a very long time scale (hours or more). There are indeed important deviations from the predictions, clearly incompatible with the spontaneous fluctuations of the detectors (see for example the real VIRGO sensitivity curve in figure 2.1). These contributions represent the other class of non-gravitational signals affecting the detector output.

2.3.2 Spurious signals from the external environment

As gws manifest themselves as a very weak mechanical effect, it is important to isolate the detectors from external, non-gravitational forces. At the same time, as the effect is finally converted to an electrical signal, no important coupling to external electromagnetic sources should exist. However, despite the technological achievements, it is practically impossible to completely decouple a detector from the external environment and regard it as an isolated system. Therefore, its output observable $x(t)$ will be also contaminated by contributions from external forces, acting independently from the effect under measurement [14]. Unlike the thermodynamical fluctuations of the instrument, however, these “disturbances” or *spurious signals* do not follow a universal model, because they are associated with different physical phenomena bearing different couplings to the detector and so

they depend on the environment and on the detector design and construction. As anticipated, they are incompatible with the description given by the fluctuation-dissipation theorem. Even if some of these effects can be modelled, the superposition of all spurious signals affecting a detector represents effectively a non-stationary process $s(t)$ whose statistics and power spectrum are very complex and poorly predictable.

It is worth listing some of the most prominent examples of spurious signals affecting current gw detectors.

- Seismic vibrations and acoustic noise, transmitted to resonant masses and interferometer mirrors through the mechanical suspension systems [72].
- The so-called Newtonian noise, i.e. the indirect sensing of nearby motion through the modulation of the static gravitational field. Ground-based interferometers are particularly affected by this problem, which represents a severe low-frequency limit (~ 10 Hz) for any gw observation with Earth-based instruments and can be defeated only with space-borne detectors like LISA [29].
- Harmonics from the power grid, coupled electrically to the detector electronics. They can be noticed in the power spectrum of the detector output as sharp spectral peaks at multiples of the fundamental frequency (50 or 60 Hz).
- Cosmic rays, interacting thermo-mechanically with the large mass of resonant detectors. They contribute to the output as short transients [51].
- Creep phenomena in the detector assembly, e.g. due to the relaxation of displacements and defects in the bulk of the suspensions. They contribute to the detector output as short bursts.
- Rings and fluctuations from the active control systems used by interferometers to operate correctly.

Many of these phenomena, especially seismic vibrations and disturbances from the power grid, are subject to anthropogenic trends, as they are due to human activities that strongly exhibit the day-night periodicity. These modulations can be indeed observed in the measured signals.

On the one hand, the practical consequence of spurious signals is that the real sensitivity curve of a detector will exhibit *excess* of noise with respect to the purely intrinsic component predicted by the fluctuation-dissipation theorem (see the right plot of figure 2.1). On the other hand, the statistic of the non-gravitational signal will be distorted from the predicted Gaussian distribution. In fact, while strong or simple spurious signals can be easily identified and removed by opportune signal monitoring, *conditioning* and *veto* procedures [52, 53] (e.g. simple band-stop filters around power grid harmonics), weak disturbances tend to confuse among the intrinsic fluctuations, effectively altering the statistic of the intrinsic noise and causing unpredictable deviations from the expected Gaussian shape.

Due to their complexity and lack of accurate models, spurious signals represent therefore a major problem for gw experiments [14]. In fact, unless the time-frequency signature of a gw signal is strong enough to be unambiguously distinguishable, it is very difficult to reliably assess the gravitational origin of any signal with a single detector. Data analysis procedures must be robust enough to recognize true gw signals and efficiently reject the spurious instrumental artifacts.

2.4 Complete output of a gw detector

Summarizing, if we assume the detector model discussed in this chapter, we can write the output signal as the superposition of gws, spurious signals and intrinsic noise,

$$\mathbf{x}(t) = s_{gw}(t) + s(t) + n(t). \quad (2.5)$$

After the deconvolution with the calibrated transfer function $G(f)$ we have

$$\begin{aligned} x(t) &= s_{gw}(t) + s(t) + n(t) = \\ &= D_{ij} h^{ij}(t) + s(t) + n(t) = \\ &= h_+(t) F^+(\vartheta, \varphi) + h_\times(t) F^\times(\vartheta, \varphi) + s(t) + n(t), \end{aligned} \quad (2.6)$$

i.e. a direct estimate of the projected gws, albeit corrupted by noise and spurious signals. Eq. 2.6 will be the basis for discussing the data analysis procedure.

Chapter 3

Data analysis in a detector network

3.1 Signal properties and data analysis requirements

To study gw bursts with the available samples $\{x\} = \{x[j]\} = \{x(jt_s)\}_{j \in \mathbb{Z}}$ of the reconstructed signal 2.6, where t_s is the sampling time, data analysis procedures have to be defined. Such procedures depend on the properties of gw bursts, noise, and spurious signals. Let us make these properties more explicit.

Gw bursts are well approximated by plane waves with an associated source direction (ϑ, φ) in the celestial sphere and two associated polarization amplitudes $h_+(t)$ and $h_\times(t)$. As discussed in chapter 1, the main properties of $h_+(t)$ and $h_\times(t)$ are the transient character and the short time scale, namely 10^1 – 10^{-3} s. As there is no control on the signal source, the burst direction, time of arrival and exact $h_+(t)$, $h_\times(t)$ waveforms are not known. Moreover, due to the source distance and intrinsic gw weakness, we expect a low signal-to-noise ratio (SNR) and a low probability of intense burst generation. Then, while $x(t)$ is dominated by the detector noise $n(t)$, bursts are immersed in the noise and *sparse*, i.e. there is a negligible probability that two strong bursts excite the detector at the same time. Finally, due to the band-limiting effect of the gw detector discussed in chapter 2, the burst waveforms can be considered “physical” (i.e. L^2) signals, with no discontinuities and with a finite signal energy

$$h_{\text{rss}} = \left[\int_{-\infty}^{+\infty} h_+^2(t) + h_\times^2(t) dt \right]^{\frac{1}{2}} \quad (3.1)$$

which should not be confused with the *physical* energy carried by the gw burst, proportional to $\int_{-\infty}^{+\infty} \dot{h}_+^2(t) + \dot{h}_\times^2(t) dt$.

The noise $n(t)$, as described in chapter 2, is a quasi-stationary, Gaussian and zero-mean random process with a given spectral power density $S_{nn}(f)$. Experience shows that, for actual gw detectors, the non-stationarity can only be noted on time scales much longer than gw bursts, so the noise is effectively stationary to a good approximation. On the contrary, Gaussianity—a

strongly desirable property—can only be achieved after the application of suitable *pre-conditioning* procedures to the data, i.e. the selection of time intervals and frequency bands where the Gaussian approximation holds to a desired extent. The price to pay is the effective reduction of duty time and frequency band available for gw searches.

In chapter 2 we also pointed out the difficulty to give accurate models for all spurious signals. Thus we must expect the presence in $s(t)$ of non-gravitational transients sharing the properties of gw bursts, notably the unknown arrival time, sparsity and finite energy. Such signals are indeed found in current detectors—either resonant [54] and interferometric [53]—and are generally called *glitches*.

The available data are therefore a set of N samples (time series) containing Gaussian noise and, occasionally, unknown transient signals of either gravitational or instrumental nature. It is common practice to split the long time series (which, possibly, lasts for months) into smaller time windows of short duration (*data sets*) to better fit computational needs and guarantee transient signal sparsity and noise stationarity in each data set. Thanks to signal sparsity, these data sets can then be analyzed independently, a feature useful, among other aspects, for parallelization. In the following, the discussion will only involve a single data set $\{x\} = \{x[j]\}_{j=1}^N$ and N will be the number of samples per data set.

Signal sparsity entails that we do not know a priori whether each data set contains a signal (either a gw burst or glitch) or only the detector noise. The first data analysis step must therefore establish the presence of a signal in the detector noise, i.e. it must perform the *detection* of candidate *events* [54].

Among the data sets identified as containing a signal, we must select the ones containing true gw burst events and reject those containing instrumental glitches. This step is the *discrimination* of gw signals, also called the *veto* of spurious events [54]. The discrimination should take place by recognizing the distinctive properties of gws, e.g. their speed of propagation, TT character and physical symmetries. As already stressed in the previous chapter, such recognition is only possible by jointly analyzing the data sets produced by different detectors, because the data set of a single instrument only carries incomplete information about the gw burst. It is worth noticing that, to maximize the efficiency of the analysis, the discrimination step should be part of the detection one. That is, an ideal gw detection algorithm should be immediately sensitive to true gws and equally insensitive both to noise and spurious signals. Such goal is clearly impossible to realize when we independently search for transient signals in the outputs of each detector.

Due to the finite SNR, the effect of the above steps can not be free from errors, and some of the selected data sets will not contain true gw bursts (*background events*). To assess both the true gw detection and the reliability of the data analysis procedure, the rate of these background events must be estimated with additional steps [54]. Usually, this involves the cancellation of the distinctive properties of gw signals, which can be done in different ways, to create an ensemble of “null” data sets. The background rate is then estimated by the simple re-application of the same data analysis to the

null data.

Once the presence of a true gw signal has been decided in a data set at a given confidence level, the final step is the estimation of the signal parameters, e.g. the source direction (ϑ, φ) and the $h_+(t)$, $h_\times(t)$ waveforms. In other words, the full reconstruction of $\mathbf{h}(t)$ can be effectively performed. As discussed in chapter 2, this is another step that only makes sense by considering the data sets from more detectors.

3.2 Signal detection as hypothesis testing

The presence of a signal alters the statistical properties of the data set, i.e. it changes its probability distribution. Therefore, detection algorithms are usually discussed in the so-called *hypothesis testing* framework [55]. Given the available data samples $\{x\}$, one formulates the *null hypothesis* \mathcal{H}_0 (no signal is present and $\{x\}$ represents noise samples) and the *alternative hypothesis* \mathcal{H}_1 ($\{x\}$ contains noise and signal). Then, the null hypothesis is either accepted or rejected by means of a suitable *decision rule*. This rule consists in checking whether the data fall into a *rejection region* $R \subset \mathbb{R}^N$, defined by means of some functional of the data set \mathcal{S} (statistics). The decision rule is then

If $\{x\} \in R$ accept \mathcal{H}_1 , otherwise if $\{x\} \notin R$ accept \mathcal{H}_0 .

\mathcal{H}_0 is a *simple* hypothesis, i.e. it specifies a unique probability distribution $P_0(\{x\})$ for $\{x\}$, e.g. the normal distribution for Gaussian noise. On the contrary, as \mathcal{H}_1 depends on the knowledge of the signal, generally it is a *composite* hypothesis, that is, it identifies under the generic name “signal” a number of different probability distributions $P_1(\{x\})$ for $\{x\}$. For instance, if we deal with signal templates, different distributions arise from different values of the template parameters.

Although \mathcal{S} may be arbitrarily complicated, it usually reduces to a simple real map $\mathcal{S} : \{x\} \rightarrow \mathbb{R}$. The rejection region is then defined by comparing $\mathcal{S}(\{x\})$ with a threshold T , namely $R_T = \{\{x\} \mid \mathcal{S}(\{x\}) \geq T\}$. In other words, the detection is equivalent to simply thresholding a real statistics of the data.

The statistical nature of the noise implies that we are unable to tell for sure whether the data contain a signal or not; instead, we can only assign a *probability* that $\{x\}$ contains only noise or noise and a signal. Therefore, the decision rule is unavoidably prone to errors. In particular, its outcome may be wrong in two cases:

- the null hypothesis is rejected when there is only noise (*false alarm*);
- the null hypothesis is accepted when there is a signal (*false dismissal*).

Conversely, the decision is correct whenever

- the null hypothesis is rejected when there is a signal (*true alarm*);
- the null hypothesis is accepted when there is only noise (*true dismissal*).

Reality Decision	Noise only	Noise + signal
Noise only	True dismissal	False dismissal
Noise + signal	False alarm	True alarm

Table 3.1: possible outcomes of a detection algorithm.

The possible outcomes are summarized in table 3.1.

Each of the four cases has an associated probability of occurrence, and one defines the false alarm probability p_{fa} (probability to decide that the signal is present when actually it is not) and the detection probability or *efficiency* p_d (probability to decide that the signal is present when it is), namely

$$\begin{aligned}
 p_{fa} &= P(\{x\} \in R_T \mid \mathcal{H}_0) \\
 p_d &= P(\{x\} \in R_T \mid \mathcal{H}_1).
 \end{aligned}
 \tag{3.2}$$

Clearly, both p_{fa} and p_d are functions of T .

The performance of the decision rule as T varies is conveniently represented by the *receiver operating characteristic* (ROC), the parametric curve defined by $(p_{fa}(T), p_d(T))$. Usually, the ROC is plotted like in figure 3.1. On the one hand, an ideal decision rule that always takes the correct de-

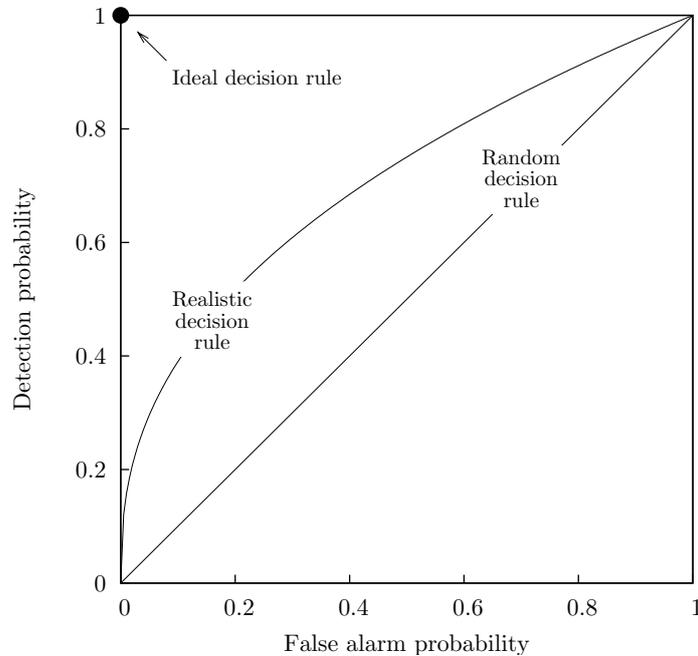


Figure 3.1: the ROC plane, showing the two extreme cases and a realistic one.

cision is represented in the ROC plane by the point $(0, 1)$. On the other hand, the most ineffective rule takes a random decision independently from the data, tracing the bisector of the ROC plane. ROC curves of realistic and useful rules lie within the triangular area delimited by these two extreme cases. ROC curves below the bisector identify rules that consistently take the wrong decision; by just inverting the logic of such rules one gets consistently correct decisions.

The desirable requirement of lowering the false alarm probability and raising the efficiency can not be attained simultaneously to an arbitrary degree. In fact, the former is obtained by raising the detection threshold T , but this unavoidably precludes the latter. This can be seen in the ROC plane by noting that the ROC curve never comes arbitrarily close to the point $(0, 1)$. Consequently, one has to pick a compromise by tuning T .

In general, however, different decision rules (or equivalently, different rejection regions or statistics) produce better or worse compromises and one looks for the “best” one for each given data set. There are several criteria to define the meaning of “best”, each with a different objective. For instance, one may be interested in robustness against noise; or, it may be best to minimize the chance of false alarms. One of the most widely used in gw searches is the well-known *Neyman-Pearson criterion*: once $p_{fa}(T)$ is determined by fixing T , choose the rejection region that maximizes $p_d(T)$. In other words, one fixes the chance of getting false alarms to an acceptable level, then looks for the rejection region that attains maximum efficiency.

If both \mathcal{H}_0 and \mathcal{H}_1 are simple hypotheses, an explicit solution to the Neyman-Pearson criterion is provided by the *Neyman-Pearson lemma*: given p_{fa} , the maximum efficiency is attained by thresholding the statistics

$$\Lambda(\{x\}) = \frac{P_1(\{x\})}{P_0(\{x\})} \quad (3.3)$$

which is the well-known *likelihood ratio*. Detection through the rejection region $R_T = \{\{x\} \mid \Lambda(\{x\}) \geq T\}$ is then called the *likelihood ratio test*. Actually, the Neyman-Pearson lemma also holds if \mathcal{H}_1 is a one-sided composite hypothesis, i.e. $P_1(\{x\}; \alpha)$ is a monotonic function of a single parameter $\alpha > 0$ or $\alpha < 0$. However, when \mathcal{H}_1 is composite and $P_1(\{x\}; \boldsymbol{\theta})$ is an arbitrary function of a parameter vector $\boldsymbol{\theta}$, the optimality of a test becomes difficult to define: there can not be a single criterion of optimality. For instance, one may want to maximize the average efficiency over the parameter space Θ , or maximize the minimum one, or one may ask for the minimum variation of the efficiency over Θ . Yet, the Neyman-Pearson lemma can be straightforwardly extended to this case by using the maximized likelihood ratio

$$\Lambda(\{x\}) = \frac{\max_{\boldsymbol{\theta} \in \Theta} [P_1(\{x\}; \boldsymbol{\theta})]}{P_0(\{x\})}. \quad (3.4)$$

$P_1(\{x\}; \boldsymbol{\theta})$, evaluated at the measured data and regarded only as a function of $\boldsymbol{\theta}$, represents the so-called *likelihood function*. The resulting test is no longer necessarily optimal, but in practice gives generally good efficiencies, and reduces to the likelihood test when the definition of “optimum” is unambiguous.

In the following, we will review a number of interesting detection algorithms and their application to gw searches.

Among the most famous detection algorithms is the *matched filter* one [55]. Suppose that the signal is known up to a scale factor A and $x[j] = Af[j] + n[j]$, where $f[j]$ is the known signal waveform and $n[j]$ is white Gaussian noise with standard deviation σ . According to the matched filter theory, A can be estimated through

$$\mathcal{A}(\{x\}) = \frac{\sum_{j=1}^N x[j]f[j]}{\sum_{j=1}^N f[j]^2}, \quad (3.5)$$

which can be shown to minimize the mean square error between the measured data $x[j]$ and the signal estimate $\mathcal{A}f[j]$. The matched filter detection rule puts the threshold on \mathcal{A} . On the one hand, \mathcal{H}_1 is a one-sided composite hypothesis, so the Neyman-Pearson lemma holds; on the other hand, the likelihood ratio can be written as $\Lambda(\{x\}) = \exp(\mathcal{A}^2/2\sigma_{\mathcal{A}}^2)$ (where $\sigma_{\mathcal{A}}^2 = \sigma^2/\sum_{j=1}^N f[j]^2$ is the variance of \mathcal{A}) and thus a threshold on \mathcal{A} is equivalent to one on Λ . In other words, thresholding \mathcal{A} is equivalent to performing the likelihood ratio test. Consequently, the matched filter detection rule is the most efficient one, for a given $p_{fa}(T)$. Another important feature is that both $p_{fa}(T)$ and $p_d(T)$ have exact expressions, which means that the ROC curve can be known analytically. Clearly, this optimum method is of poor practical utility to detect gws, because the signal waveform is not known. Still, it is useful as an asymptotic limit on ROC curves of realistic, less efficient gw detection algorithms.

A generalization of the matched filter detection rule allows one to lack the knowledge about the signal arrival time. This immediately turns \mathcal{H}_1 into a composite hypothesis and the Neyman-Pearson lemma is no longer applicable. One can then resort to eq. 3.4 and get, in general, good efficiencies, though the algorithm false alarm probability and efficiency can only be estimated by means of Monte Carlo simulations.

A further matched filter generalization relaxes the required knowledge of the full signal waveform. The test must be performed by using eq. 3.4 and maximizing over a bank of waveforms $\{f_{\theta}[j]\}_{\theta \in \Theta}$ (templates), where the parameter vector θ labels each waveform of the bank. This test is used to look for gw signals that can be given exact waveforms; the typical example are inspiralling compact binaries.

To detect transient signals without templates, another efficient algorithm is the *power filter*. Suppose that we know how to localize the power of the transient signal in the time-frequency plane by convolving $x[j]$ with a suitable filter $q[j]$. Then, the *local power* $E_q = \sum_{j=1}^N (q * x)[j]^2$ can be shown to be an optimal statistics, a threshold on E_q being equivalent to the likelihood test [56]. In practice, especially in the gw case, the signal time-frequency support is poorly known and one does not know how to build the localizing filter. In this case the filter is omitted and instead the *total power* $E = \sum_{j=1}^N x[j]^2$ can be used as the statistics. Optimality is lost and the resulting efficiency is rather low, but the method drops any assumption about the signal waveform, at least as long as it is fully included in the

Reality Decision	Noise only	Noise + glitch	Noise + gw
Noise only	True dismissal	True dismissal	False dismissal
Noise + glitch	True dismissal	True dismissal	False dismissal
Noise + gw	False alarm	False alarm	True alarm

Table 3.2: possible outcomes of a genuine gw detection algorithm, assuming signal sparsity. “Alarm” is to be intended with respect to gw detection only.

data set. For this reason it makes little sense to use less efficient detection algorithms, and the total power test becomes interesting as a lower bound to ROC curves of other statistics.

The search for powerful tests should therefore focus on the region of the ROC plane between the ROC of the matched filter and the one of the total power.

3.3 Discrimination and multiple hypotheses test

In the gw case, the spurious signals represent an effective additional source of transients in the data sets, and the general two-hypotheses framework discussed in the previous section must be extended to tackle this problem. We identified two approaches.

A straightforward way is simply to treat spurious transients as belonging to \mathcal{H}_0 (e.g. they are part of the detector noise), thus maintaining the structure of the hypothesis test. Then \mathcal{H}_0 becomes a composite hypothesis on its own. Unfortunately, the identification of efficient detection rules becomes a difficult task, as neither the probability density associated with \mathcal{H}_0 (a superposition of noise and poorly-modeled spurious signals) nor the one associated with \mathcal{H}_1 (determined by gws) are known accurately.

Another, more general, approach is to extend table 3.1 to take into account spurious signals. By invoking signal sparsity, each data set may contain either a gw burst, a glitch or only noise; consequently, \mathcal{H}_1 can be split into a “gw” alternative hypothesis \mathcal{H}_{gw} and a “spurious” one \mathcal{H}_s (table 3.2). The problem is then complicated to a *multiple* hypothesis test [55], which may involve more than a single detection threshold and whose performance can not be characterized by a single ROC curve. Nevertheless, thanks to signal sparsity, one may separately compute “gw” and “glitch” ROC curves, representing the performance of gw and glitch detection in the noise, or a ROC curve may be computed for distinguishing \mathcal{H}_{gw} from \mathcal{H}_s .

Ideally, the analysis algorithm should be able to efficiently distinguish among the three hypotheses. This would allow one to detect gws and, at the same time, to monitor the network instrumental glitches, achieving the most complete information about both the physical phenomenon under study and the measurement instrument represented by the network.

3.4 Networks of gw detectors

With a network of $M > 1$ detectors, the available data set is extended from a scalar time series $\{x\}$ ($N \times 1$ samples) to M time series, conveniently represented as a discrete M -vector $\{\boldsymbol{x}\}$ ($N \times M$ samples). We identify a general data analysis procedure for accomplishing the goals discussed in the previous sections. It is sketched in figure 3.2, where we have also included the preliminary calibration step for completeness.

The basic idea behind such analysis pipeline is to overcome our inability in recognizing the gw properties in the single outputs by constructing a number of synthetic *virtual channels* (or data sets) out of $\{\boldsymbol{x}\}$. These channels are then used to separately accomplish each goal of section 1, i.e. detection, discrimination and parameter estimation. For this reason, they are grouped as “detection”, “discrimination” and “estimation” channels, though they may be many more streams of information.

The pre-conditioning step represents the tasks performed locally to each detector to validate the assumptions made about the noise, e.g. Gaussianity. As anticipated, these may include the removal of time epochs and frequency bands where the detector output $x(t)$ is not suitable for gw searches, for instance because it is dominated by the spurious contribution $s(t)$, thus bearing a complex and unpredictable behaviour. Pre-conditioning may also include whitening procedures to remove the noise correlation. Nevertheless, it must be stressed that pre-conditioning tasks performed locally to each detector are likely to impair the further analysis. For instance, removing

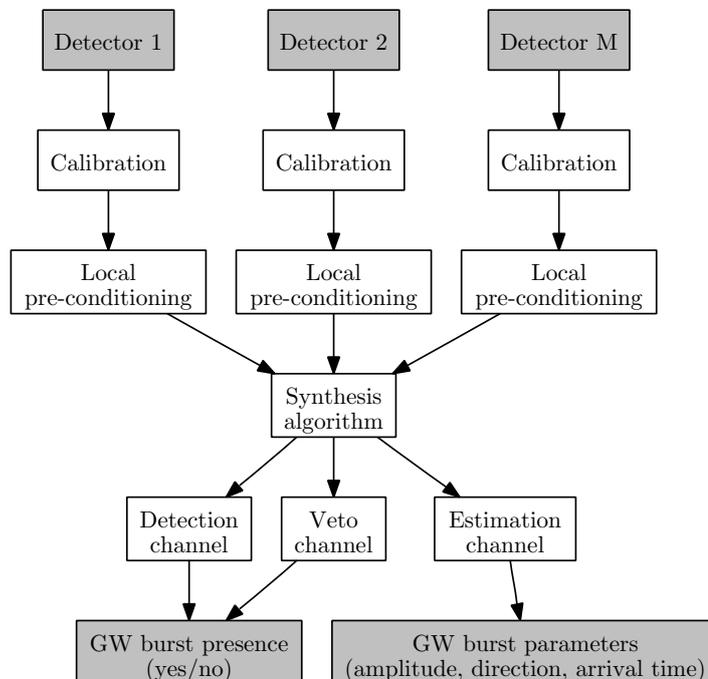


Figure 3.2: schematic of a data analysis pipeline for identifying gw bursts with a network of detectors.

time epochs and frequency bands effectively reduces the data available for the analysis, unavoidably reducing the chance of detecting genuine gws. Any operation performed locally should be avoided as much as possible or delayed at later stages in the pipeline, i.e. within the synthesis stage itself. In fact, this allows one to carry out more targeted conditioning, taking into account the physical properties of gws which can only be appreciated at the synthesis stage.

The synthesis algorithm is the set of operations that define the synthetic output channels and it is the part of the pipeline where the optimization should concentrate. The detection channel contains information that allows one to distinguish \mathcal{H}_0 from \mathcal{H}_{gw} and \mathcal{H}_s . For example, it may represent the instantaneous energy of the network sample $\{\mathbf{x}\}$, to perform a threshold detection similar to the total power test. The discrimination channel should gather information distinguishing \mathcal{H}_{gw} from \mathcal{H}_s . For instance, it could contain the trace of the gw tensor $\text{tr}[\mathbf{h}(t)]$, which must be 0 assuming general relativity. A threshold crossing on such a channel would then rule out \mathcal{H}_{gw} and mark (veto) the data set as contaminated by a glitch. Together, the detection and estimation channels assess the presence of a genuine gw burst in each data set, telling \mathcal{H}_{gw} from either \mathcal{H}_0 or \mathcal{H}_s . Finally, if this is the case, the estimation channel reconstructs the parameters of $\mathbf{h}(t)$ or, equivalently, its time-dependent components.

To assess the effectiveness of the full analysis pipeline, background estimation can then be accomplished by creating the ensemble of null data sets and by running the same pipeline on them. Thanks to the distinctive properties of gws, the null ensemble can be easily produced in different ways. An usual procedure is shifting the M detector outputs by different time delays, larger than the time taken by gws to travel the Earth diameter (which is $2\tau_E = 2r_E/c \simeq 4 \cdot 10^{-2}$ s, where $r_E \simeq 6 \cdot 10^6$ m is the mean Earth radius). The phase information of any possible gravitational signal is then destroyed, as physically gws can only contribute to each detector within a time window not larger than $2\tau_E$. Another approach is the so-called method of *surrogate data*: the removal of any phase coherence takes place by transforming the data sets to the frequency domain ($\{x\} \rightarrow \{X\}$), then randomizing the phase spectra—by setting $\{\arg(X)\}$ to M random numbers in the range $[0; 2\pi]$ —and finally transforming back to the time domain ($\{X\} \rightarrow \{x\}$). Evidently, the method of surrogate data has the property of exactly conserving the power spectra of the data sets.

A desirable property of an analysis pipeline is the ability to estimate the background for each data set, e.g. by creating the ensemble of surrogate data out of each single data set. In other words, the analysis should not need data in a much longer time interval than the duration of a single data set. This feature is important for assessing the performance of the network and possibly to study its time dependence. In fact, it represents an extremely powerful proof of the robustness against non-gravitational spurious disturbances with annual, monthly or even daily periodicities.

Calibration errors in the transfer functions of the detectors, both in phase and amplitude, clearly limit the effectiveness of any analysis pipeline, e.g. because they tend to disrupt the physical signatures of gws. Therefore, it

is also important to characterize and improve the robustness of the analysis against such errors.

We now briefly review the burst data analysis techniques currently being used or investigated.

Data produced by the resonant-bar observatory is being analyzed by the International Gravitational Event Collaboration (IGEC) [22, 23]. The analysis procedure first performs individual detection on the single-detector data sets, in the framework of hypothesis testing discussed above. Single-detector alarms are defined through a modified matched filter decision rule, where the filter is matched to the known, narrow-band detector impulse response $g(t)$. This statistics performs well for δ -like gw bursts with a constant Fourier transform over the detector bandwidth. Individual event lists then undergo a *coincidence search*, looking for events occurring in more detectors within short time windows. The rate of such coincidences is estimated. The background coincidence rate is then estimated by rigidly shifting single-detector event lists by unphysical delays (i.e. larger than $2\tau_E$) and repeating the coincidence search. An excess of coincidence rate in the non-shifted event lists can be associated with gw bursts. The IGEC analysis does not actually fit the scheme in figure 3.2, as there is no synthesis of virtual channels and only partial information about the gw parameters can be reconstructed, namely the burst amplitude and time of arrival.

Indeed, such a method has a number of drawbacks. It requires a network of aligned detectors, because they must receive an identical gw signal $s_{gw}(t)$. Then, as the detection process is performed independently on each instrument, the global detection efficiency can be written as $\varepsilon_{\text{net}} = \prod_{m=1}^M \varepsilon_m$, where ε_m is the detection efficiency for detector m . $\varepsilon_m < 1$ implies $\varepsilon_{\text{net}} < \min(\varepsilon_m)$, i.e. the global efficiency is bounded from above by the *less* efficient detector. Furthermore, this method exploits neither the physical gw properties nor the phase coherence of gw signals—for this reason it is called an *incoherent* search—but it rather looks for coincident excitations of similar instruments. Thus, its discrimination power is essentially based on the hypothesis that occurrences of spurious signals in one detector are uncorrelated with those in the other detectors. This is a limitation because, as discussed later, the network delivers much more information.

Interferometers have a large bandwidth with respect to resonant bars. Hence, detection of unknown signals through filters matched to the instrument impulse response is no longer effective. However, the coincidence search depicted above can be extended to networks of interferometers by means of time-frequency decompositions and sub-optimal power filters (or similar statistics with minimal assumptions about the signal) [57, 58]. One of the main implementations of such a search is the WaveBurst algorithm [59]. It works by first finding clusters of high-energy coefficients in a time-frequency decomposition of the data sets, depending on a preliminary threshold. Then, it looks for clusters in the different time-frequency maps that bear consistency within a suitable time window. Finally, a cluster *significance* is evaluated and compared to a second threshold to attain the final event detection. Clearly this technique still represents an incoherent search and thus carries similar drawbacks, notably the requirement of aligned antenna

patterns.

In the late years, the interest of the gw community has been shifting towards *coherent* burst analysis methods. The name comes from the fact that such methods exploit the (expected) strong phase coherence of the gw waveform in order to recognize the presence of the burst in the outputs of the M detectors. The M data sets are merged into one or more statistics where the detection is then performed, rather than carrying out M independent tests and then combining the outcomes with logical operations. Thus, the efficiency limit of coincidence-like incoherent searches should be avoided. Moreover, coherent methods can be designed to take into account the differently aligned detector antenna patterns. This allows one to use coherent methods with networks of arbitrarily aligned detectors. Finally, only the coherent combination of the data sets can estimate the full signal parameters. Therefore, a coherent analysis is a mandatory part of figure 3.2.

A straightforward procedure to enforce the phase coherence of the signal is a direct cross-correlation between the detector outputs. The cross-correlation allows one to build powerful statistics, such as the r -statistic, to reject non-gravitational events produced by noise and glitches [62] and also makes it possible to localize the gw source in the sky [63].

Coherent methods can also be constructed by applying the maximum-likelihood principle to the network response, thus constructing a “network likelihood functional” that intrinsically takes into account the cross-correlation between different detectors [64]. Such methods can reconstruct the source direction as well. Although they can be used both with aligned and misaligned detectors, it has been demonstrated that in the latter case they are very sensitive to an issue known as *two-detector paradox*. This problem arises from a subtle degeneracy in the detector network, that we will discuss later, and tends to produce unphysical solutions that have to be carefully excluded by *regularizing* the problem. A practical implementation of maximum-likelihood coherent search is an extension of the WaveBurst algorithm that realizes a likelihood-based coherent detection in the time-frequency domain [65].

Coherent analyses have also been defined by means of constructing suitable linear combinations of the outputs, that either estimate or cancel the gw polarization amplitudes. On the one hand, the former combinations can be used for the reconstruction of the signal waveforms h_+ and h_\times [60]. On the other hand, those that cancel the gw contribution, commonly called *null streams*, can be used to infer the source direction [60] and at the same time to efficiently reject instrumental glitches [61]. This kind of coherent analysis best matches the scheme in figure 3.2. It is promising because it directly tackles the so-called *gw inverse problem* that will be discussed in the rest of the section.

In order to discuss realistic detector networks, we report in table 3.3 the relevant geometric parameters of the currently operating interferometers, i.e. their geographical location and arm orientation [66]. The locations are given in geocentric coordinates, where the origin is the center of Earth, the z axis points towards the North pole, the x axis is directed towards the Greenwich meridian and the y axis is so that the coordinate system is right-

Detector	ϑ	φ	α_1	α_2
VIRGO	46.3333	10.50	70.5674	160.5674
LIGO Livingston	59.437	269.266	197.7165	287.7165
LIGO Hanford	43.545	240.592	125.9994	215.9994
GEO600	37.7549	9.8072	21.6117	115.6117
TAMA300	54.3264	139.5392	180.	270.

Table 3.3: geometric parameters of the currently operating interferometers. (ϑ, φ) is the detector location in geocentric spherical coordinates. α_1 and α_2 are the orientations of the two arms with respect to the local East. All values are in degrees.

handed. Spherical coordinates are used, where $0^\circ \leq \vartheta \leq 180^\circ$ is the latitude (0° at the North pole, 180° at the South pole) and $0^\circ \leq \varphi < 360^\circ$ is the longitude (0° is the Greenwich meridian). To provide a first-glance look at the locations and relative orientation of the detectors, we report in figure 3.3 their scattering on the Earth surface and their polarization independent antenna patterns $\widehat{F}(\vartheta, \varphi)$. It is worth noticing the similar orientation of the LIGO instruments, which allows one to perform incoherent burst searches in their data sets. The tiny misalignment of the two detectors is an unavoidable effect of the Earth curvature.

3.4.1 Network response to gws

According to eq. 2.6, the outputs of an M -detector network can be written in the time domain as

$$\begin{aligned}
x_m(t) = & h_+(t + \tau_m(\vartheta, \varphi)) F_m^+(\vartheta, \varphi) + \\
& + h_\times(t + \tau_m(\vartheta, \varphi)) F_m^\times(\vartheta, \varphi) + \\
& + s_m(t) + n_m(t),
\end{aligned} \tag{3.6}$$

where the index $m = 1 \dots M$ labels each detector. The time delays $\tau_m(\vartheta, \varphi)$ account for the different detector positions on the Earth surface, which lead to different burst arrival times on each detector. Setting the origin at the center of the Earth and assuming that gws propagate at speed c ,

$$\tau_m(\vartheta, \varphi) = -\frac{\mathbf{r}_m \cdot \mathbf{k}(\vartheta, \varphi)}{c}, \tag{3.7}$$

where \mathbf{r}_m is the position vector of detector m and $\mathbf{k}(\vartheta, \varphi)$ is the wave vector given by eq. 1.12. As an example, we sketch in figure 3.4 the simulated response of the LIGO-VIRGO network to a sample gw burst.

Eq. 3.6 can be recast in vector/matrix form as

$$\mathbf{x}(t) = \mathbf{F}(\vartheta, \varphi) \mathbf{h}(t) + \mathbf{s}(t) + \mathbf{n}(t), \tag{3.8}$$

where

$$\mathbf{h}(t) = \begin{bmatrix} h_+(t) \\ h_\times(t) \end{bmatrix}, \quad \mathbf{s}(t) = \begin{bmatrix} s_1(t) \\ \vdots \\ s_M(t) \end{bmatrix}, \quad \mathbf{n}(t) = \begin{bmatrix} n_1(t) \\ \vdots \\ n_M(t) \end{bmatrix}. \tag{3.9}$$

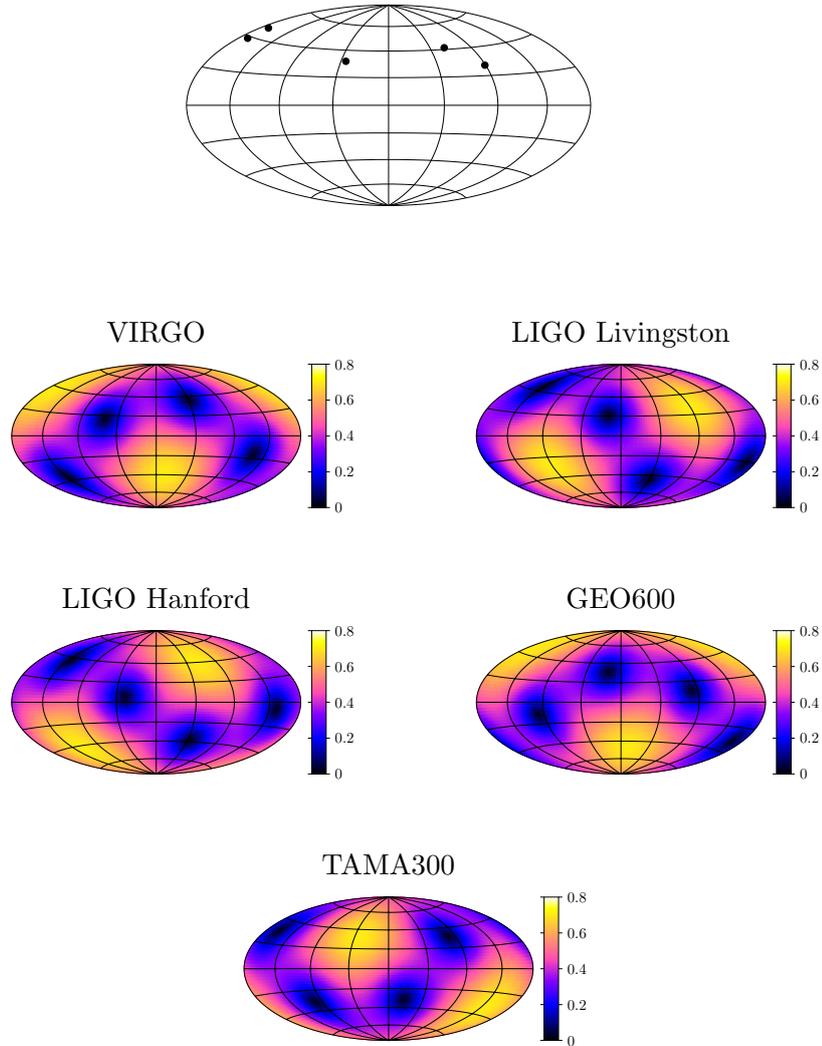


Figure 3.3: on the top, location of the interferometers on the Earth surface. From left to right we find VIRGO, GEO600, TAMA300, LIGO Hanford and LIGO Livingston. Below, sky maps of their polarization-averaged antenna patterns $\hat{F}(\vartheta, \varphi)$. Note the characteristic four zeroes of each interferometer and the similar orientation of the two LIGO instruments.

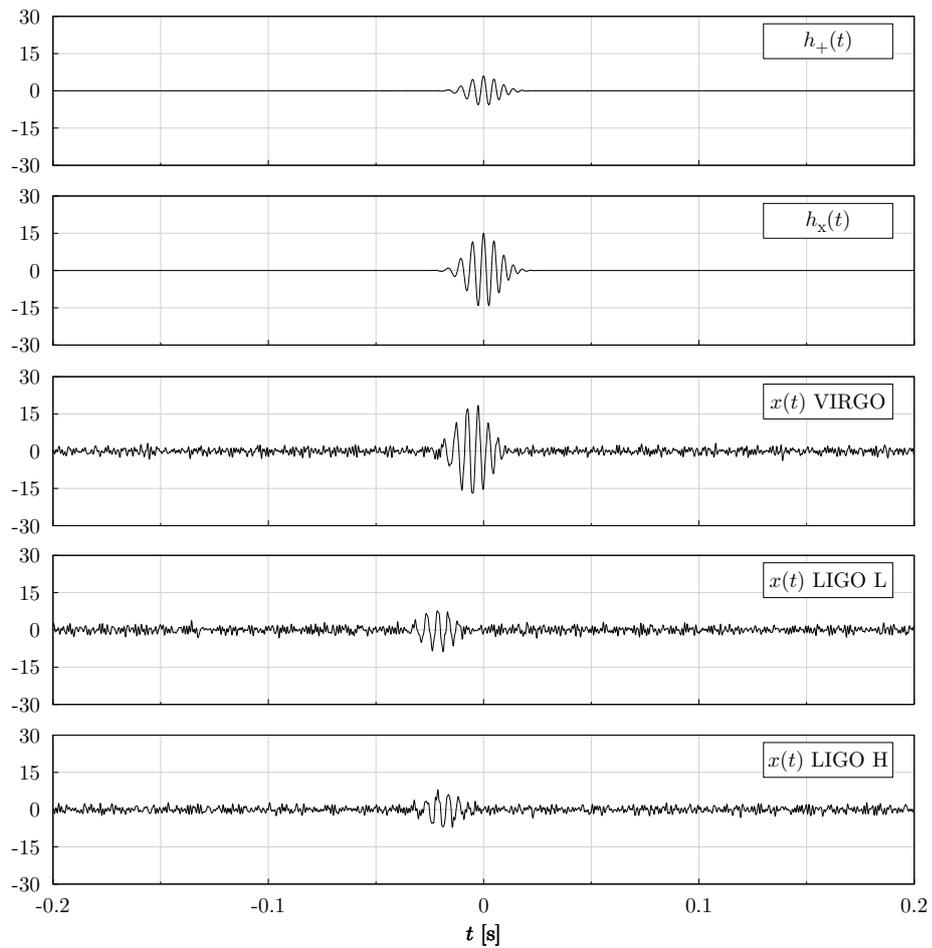


Figure 3.4: simulated response of the LIGO-VIRGO network to a sample gw burst. The burst is a linearly polarized sine-Gaussian transient, coming from the direction $(\vartheta, \varphi) = (59.7^\circ, 270^\circ)$. The two upper plots show the true burst waveforms. The lower plots show the outputs of the interferometers, contaminated by white Gaussian noise with $\sigma = 1$. Note the different amplitudes and arrival times.

Note that with such notation $\mathbf{h}(t)$ no longer represents the full gw *tensor*, but merely an artificial *2-vector* constructed with the polarization amplitudes. We will make no further reference to the gw tensor as $\mathbf{h}(t)$ to avoid confusion. The $M \times 2$ matrix

$$\mathbf{F}(\vartheta, \varphi) = \begin{bmatrix} T_1(\vartheta, \varphi) & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & T_M(\vartheta, \varphi) \end{bmatrix} \begin{bmatrix} F_1^+(\vartheta, \varphi) & F_1^\times(\vartheta, \varphi) \\ \vdots & \vdots \\ F_M^+(\vartheta, \varphi) & F_M^\times(\vartheta, \varphi) \end{bmatrix} \quad (3.10)$$

is known in literature as the *network response matrix* [68]. The first, diagonal factor describes the time delays by means of the time-shift operators $T_m(\vartheta, \varphi)$. Thus, it accounts for the gw propagation speed and the geographical distribution of the detectors. The second factor contains the antenna patterns and so it encodes the angular sensitivity of the network. Equivalently, extending the antenna pattern discussion of chapter 1, it represents a manifestation of the tensor and TT characters of gws.

Such vector form of the output is interesting because it factorizes many aspects of the network into the response matrix, underlining the linear, algebraic nature of the problem. The whole network is regarded like a single “virtual instrument” that responds to gws as a linear system. Moreover, two crucial properties of gws, namely the propagation speed and tensor/TT character, give rise to factorized contributions in the network response. This suggests that the corresponding tests can be performed independently.

However, an important aspect of the network sensitivity is left in $\mathbf{n}(t)$: the characteristics of the detector noise processes. In principle this information can be brought back into \mathbf{F} with a “network whitening” procedure. In fact, assuming that the detector noises are stationary, so that they are well described by their sensitivity curves and can be independently whitened by M whitening operators W_m , one can recast the network output as $\tilde{\mathbf{x}}(t) = \mathbf{W}\mathbf{x}(t) = \tilde{\mathbf{F}}\mathbf{h}(t) + \tilde{\mathbf{s}}(t) + \tilde{\mathbf{n}}(t)$, with

$$\begin{aligned} \tilde{\mathbf{F}} &= \begin{bmatrix} W_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & W_M \end{bmatrix} \begin{bmatrix} T_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & T_M \end{bmatrix} \begin{bmatrix} F_1^+ & F_1^\times \\ \vdots & \vdots \\ F_M^+ & F_M^\times \end{bmatrix} \\ \tilde{\mathbf{s}}(t) &= \mathbf{W}\mathbf{s}(t) \\ \tilde{\mathbf{n}}(t) &= \mathbf{W}\mathbf{n}(t). \end{aligned} \quad (3.11)$$

By definition, $\tilde{\mathbf{n}}(t)$ now represents M Gaussian, white processes with unit variance and thus it no longer carries information about network characteristics. On the other hand, the new response matrix carries the information about the frequency sensitivity of the network, as well as the angular sensitivity and geographical distribution. The generalized network response matrix takes a simple form in the frequency domain, where the time-shift operators $T_m(\vartheta, \varphi)$ are purely phase factors and the noise whitening can be

carried out by a normalization to the sensitivity curves:

$$\tilde{\mathbf{F}}(\vartheta, \varphi, f) = \begin{bmatrix} \frac{\exp[i 2\pi f \tau_1(\vartheta, \varphi)]}{\sqrt{S_{hh,1}(f)}} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \frac{\exp[i 2\pi f \tau_M(\vartheta, \varphi)]}{\sqrt{S_{hh,M}(f)}} \end{bmatrix} \begin{bmatrix} F_1^+(\vartheta, \varphi) & F_1^\times(\vartheta, \varphi) \\ \vdots & \vdots \\ F_M^+(\vartheta, \varphi) & F_M^\times(\vartheta, \varphi) \end{bmatrix}. \quad (3.12)$$

This form allows one to easily understand the consequence of including detectors of different sensitivity in the network. The response of a noisy detector is automatically suppressed by the first factor and vanishes in the limit of arbitrarily large noise variance. In other words, noisy detectors are effectively removed from the network. This justifies the intuitive idea that including a very noisy detector in a network brings little contribution to the solution of the problem.

As a further generalization, one may even account for slow noise non-stationarities by using an intermediate time-frequency domain rather than the pure frequency domain, i.e. by fully specifying the complete time-frequency behaviour of the sensitivity curves. This would take into account, for instance, an increased noise variance in some frequency bands occurring during daytime. The resulting response matrix would then bear a complex dependency—sky direction (ϑ, φ) , time and frequency—which would reflect the truly complex behaviour of the virtual instrument represented by the network.

In the following we will omit the dependencies of \mathbf{F} and we will also avoid specifying if $\tilde{\mathbf{F}}$ or \mathbf{F} is being used, when such information is of secondary importance.

The formalism introduced above suggests to define a polarization independent “network sensitivity” as the matrix norm¹ $\|\mathbf{F}\|$. As a function of (ϑ, φ) , this figure of merit plays the role that $\hat{F}(\vartheta, \varphi)$ plays for a single detector, i.e. it represents the network sensitivity to an ensemble of gw sources with uniformly distributed polarization angles.

Intuitively, one expects that scattering the detectors on the Earth surface should not affect the overall sensitivity of the network. Indeed, it can be verified numerically that $\|\mathbf{F}\|$ does not depend on the factor containing the time delays. This can be also seen from 3.12: the delays are just phase factors that can not alter norms. Therefore, networks of co-located or geographically distributed instruments have the same sensitivity (provided the detectors retain the same orientation) although, as discussed later, the propagation speed can only be tested by the latter setup. This stresses even more the separation between the effect of the finite gw propagation speed and the tensor/TT properties.

3.4.2 Inverse problem and network conditioning

Formally, eq. 3.8 represents a well-known *inverse problem* in presence of noise: under the assumption that a gw is coming from a *single* source located at (ϑ, φ) , we want to estimate the unknown amplitudes $h_+(t)$, $h_\times(t)$ from

¹We use the induced 2-norm—the largest singular value of \mathbf{F} .

the measured data $\mathbf{x}(t)$; we are neglecting the spurious contribution for the moment. Inverse problems are common in science and engineering [67] and they can be attacked with a vast mathematical framework. The discussion of the gw inverse problem is rather recent [60] and the best formalism for its solution seems to be the Moore-Penrose inversion [68].

For 2-detector networks \mathbf{F} is a simple 2×2 matrix and the inverse problem can be immediately solved through the standard matrix inversion. However, when $M > 2$, the problem is over-determined because we have 2 unknowns and M equations. As $M - 2$ equations represent “noisy” constraints, in general an exact solution for $h_+(t)$, $h_\times(t)$ does not exist. Accordingly, the network response is rectangular and can not be straightforwardly inverted. Nevertheless, there are many generalization to matrix inversion that also work for rectangular matrices, collectively known as *generalized inverses* or *pseudo-inverses* [69]. They correspond to different ways to look for approximate solutions (following different criteria) and thus they are useful in solving the overdetermined gw inverse problem. The Moore-Penrose inversion defines an approximate solution through the minimization of the least-squares functional

$$L[\mathbf{h}] = \|\mathbf{x} - \mathbf{F}\mathbf{h}\|^2 \quad (3.13)$$

where the Euclidean vector 2-norm is used. This functional has the property of resulting from a maximum-likelihood approach. Its minimization entails the search for a 2-vector \mathbf{h} that shortens the Euclidean distance between the measured data and the network response to a gw with amplitudes given by \mathbf{h} itself. Variation of $L[\mathbf{h}]$ with respect to h_+ and h_\times leads to the *normal equations* $\mathbf{M}\mathbf{h} = \mathbf{F}^T\mathbf{x}$, where $\mathbf{M} = \mathbf{F}^T\mathbf{F}$ is a 2×2 square matrix that can be readily inverted to give the solution

$$\mathbf{h}(t) = \mathbf{F}^\dagger\mathbf{x}(t). \quad (3.14)$$

The $2 \times M$ matrix

$$\mathbf{F}^\dagger = \mathbf{M}^{-1}\mathbf{F}^T = (\mathbf{F}^T\mathbf{F})^{-1}\mathbf{F}^T \quad (3.15)$$

is known as the *Moore-Penrose inverse* of the network response matrix. It has the desirable property of reducing to the standard matrix inverse for 2-detector networks.

Notably, the Moore-Penrose inversion provides an *unbiased* estimate or “reconstruction” of the gw amplitude: for each time instant, the expectation value over noise realizations $\langle \mathbf{F}^\dagger\mathbf{x}(t) \rangle$ equals $\mathbf{h}(t)$, provided that \mathbf{F}^\dagger is evaluated at the correct source location (ϑ, φ) (figure 3.5). Then, the estimates given by $\mathbf{F}^\dagger\mathbf{x}(t)$ can be regarded as synthetic “output channels” (two for each source location) of the virtual instrument represented by the network. It must be noted that the impulsive character of the gw signal is never assumed in the formulation of the inverse problem: as a result, the formalism works not only for gw bursts, but actually for gws of any kind. The only requirement is that we have at most one point-like gw source for each data set.

Unfortunately, however, for an arbitrary network the inverse problem expressed by 3.8 falls into the family of *ill-conditioned* problems [68, 70].

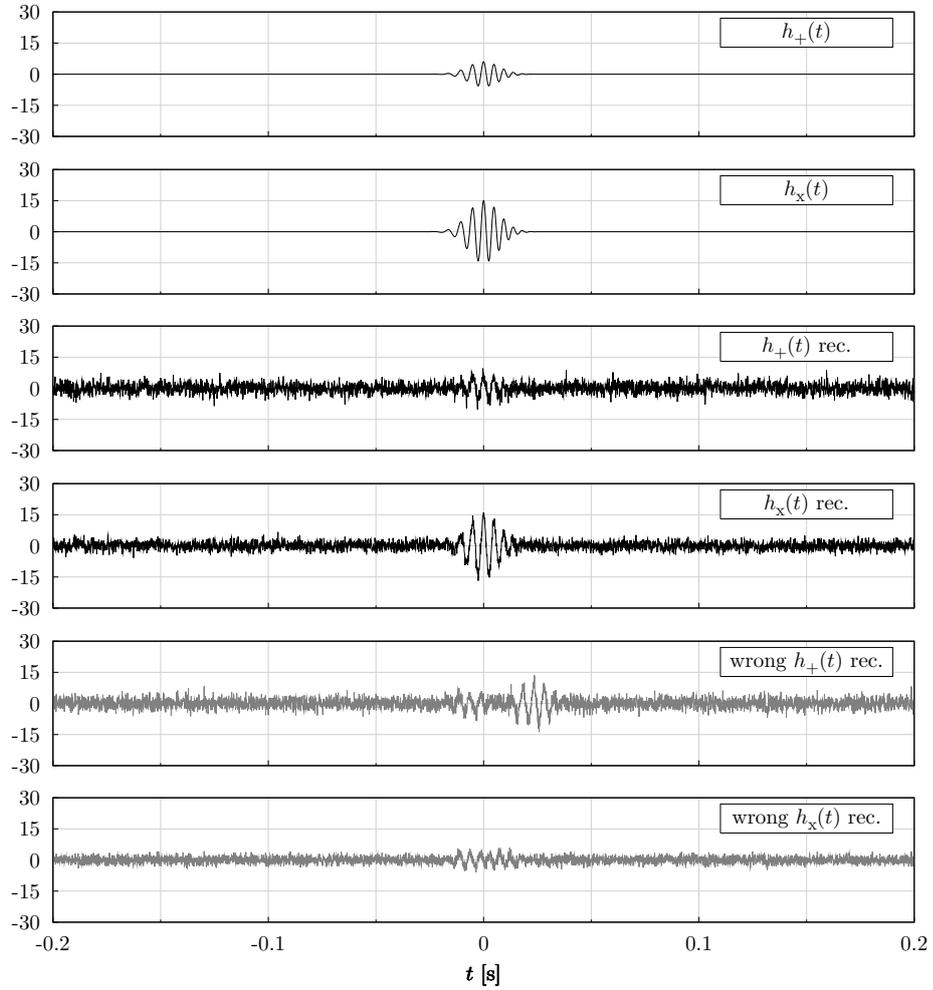


Figure 3.5: reconstruction of a simulated gw burst with the LIGO-VIRGO network by application of the Moore-Penrose inverse \mathbf{F}^\dagger to the data samples. The first two plots show the true burst waveforms—a linearly polarized sine-Gaussian transient coming from the direction $(\vartheta, \varphi) = (66^\circ, 137^\circ)$. The two middle plots are the waveforms estimated from the noisy data, with \mathbf{F}^\dagger evaluated at the exact source direction. The shaded plots are the same but using a very different direction. Note the disruption of the gw waveforms caused by the de-synchronization of the time delays.

Such problems are characterized by a strong instability with respect to small variations of the available data, e.g. they are extremely sensitive to noise or rounding errors. In the gw case, for some locations of the celestial sphere the network response matrix may suffer a subtle singularity known as *rank deficiency*. The question is that, for those values of (ϑ, φ) , the range vectors of \mathbf{F} , represented by its columns

$$\mathbf{F}^+ = \begin{bmatrix} F_1^+ \\ \vdots \\ F_M^+ \end{bmatrix} \quad \mathbf{F}^\times = \begin{bmatrix} F_1^\times \\ \vdots \\ F_M^\times \end{bmatrix}, \quad (3.16)$$

become nearly proportional. This implies that \mathbf{F} , whose nominal rank is 2, is very close to a rank-1 matrix and thus its inversion leads to unstable solutions: small perturbations in $\mathbf{x}(t)$ are greatly amplified by \mathbf{F}^\dagger . The singularity can be made more explicit by noting that \mathbf{M} becomes nearly singular as well, as it can be shown that one of its two eigenvalues approaches 0 while the other does not. Looking at the $L[\mathbf{h}]$ minimization, the problem is that the network effectively does *not* respond to one of the components of \mathbf{h} : therefore, the distance is very small for a wide set of 2-vectors and looking for the minimizing one becomes ambiguous.

Ill-conditioning is a major problem which, in the gw case, is at the origin of a number of pathologies and singularities (like the two detector paradox [64]) that also affect other data analysis strategies. The practical consequence is that the gw estimate given by 3.14, though still unbiased, is swamped by the noise fluctuations, whose variance gets strongly amplified—even by orders of magnitude—with respect to the original noise of the detectors (figure 3.6). More precisely, it can be shown numerically that *one* gw polarization amplitude can be estimated with reasonable noise contamination, while the other is blinded by a much large noise variance. This is in agreement with the above consideration that the network is not sensitive to one of the \mathbf{h} components.

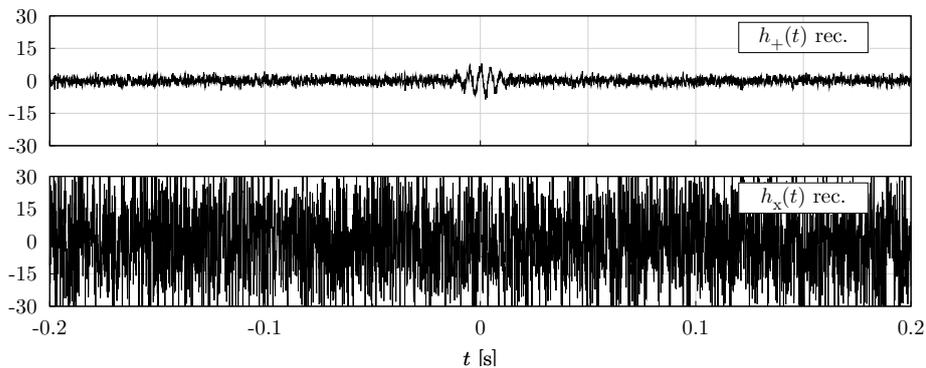


Figure 3.6: failure in reconstructing the “cross” polarization of a simulated gw burst that is coming from a direction with a strongly ill-conditioned response matrix. The $h_\times(t)$ estimate is totally blinded by the amplified noise fluctuations.

The degree of ill-conditioning of a matrix \mathbf{A} can be conveniently described by the *condition number*

$$\text{cond}(\mathbf{A}) = \|\mathbf{A}\| \|\mathbf{A}^\dagger\|. \quad (3.17)$$

Perfectly invertible matrices have $\text{cond}(\mathbf{A}) = 1$, while matrices close to singularity have $\text{cond}(\mathbf{A}) \gg 1$. Therefore, a sky map of $\text{cond}[\mathbf{F}(\vartheta, \varphi)]$ represents another network “figure of merit”, together with the norm of the network response matrix $\|\mathbf{F}(\vartheta, \varphi)\|$ introduced in the previous section. Regions in which $\text{cond}(\mathbf{F}) \gg 1$ are characterized by diverging noise variance in the network output channels given by 3.14.

Several techniques have been developed to address rank deficiency and the consequent ill-conditioning of an inverse problem. One of the most widely used is the *Tikhonov regularization* [71], which has been applied in the gw case [68]. Its essence is the modification of the least-squares functional 3.13 with the addition of a “regularizing” functional $\Omega[\mathbf{h}]$ controlled by a parameter $\gamma \geq 0$,

$$L_\gamma[\mathbf{h}] = \|\mathbf{x} - \mathbf{F}\mathbf{h}\|^2 + \gamma\Omega[\mathbf{h}]. \quad (3.18)$$

$\Omega[\mathbf{h}]$ and its strength γ are chosen so that the resulting Moore-Penrose inverse no longer suffers from rank deficiency. In the gw case a quadratic form has been used for the regulator, i.e. $\Omega[\mathbf{h}] = \mathbf{h}^T \mathbf{\Omega} \mathbf{h}$ with $\mathbf{\Omega}$ a suitable 2×2 matrix [68]. $\mathbf{\Omega}$ can be made a function of (ϑ, φ) , adapting it to the direction-dependent degree of ill-conditioning. The construction of the regularized pseudoinverse then uses $\mathbf{M}_\gamma = \mathbf{M} + \gamma\mathbf{\Omega}$ instead of the original \mathbf{M} . The role of $\mathbf{\Omega}$ is essentially to compensate the vanishing eigenvalue of \mathbf{M} . The resulting Moore-Penrose inverse reads

$$\mathbf{F}_\gamma^\dagger = \mathbf{M}_\gamma^{-1} \mathbf{F}^T = (\mathbf{F}^T \mathbf{F} + \gamma\mathbf{\Omega})^{-1} \mathbf{F}^T. \quad (3.19)$$

Remarkably, $\mathbf{\Omega}$ and γ can be chosen to attain $\text{cond}_\gamma(\mathbf{F}) = \|\mathbf{F}\| \|\mathbf{F}_\gamma^\dagger\| = 1$ everywhere, thus stabilizing the estimation of $h_+(t)$ and $h_\times(t)$ with respect to fluctuations in $\mathbf{x}(t)$. The practical result is that the noise variance is no longer amplified.

Nevertheless, such correction is not for free. In fact, every regularization procedure, by its own nature, introduces a bias in the estimation: this is the price to pay for a stable solution to the ill-conditioned inverse problem. As a consequence, the regularization strength γ must be tuned accurately, reaching a compromise between instability due to noise and bias due to regularization.

3.4.3 Actual and ideal networks

The definitions of two network figures of merit, global sensitivity and condition number, immediately suggest a number of desirable properties in a network.

1. The absence of blind directions: the regions in the celestial sphere for which $\|\mathbf{F}\| \simeq 0$ should be small or null.

2. The good conditioning of the response matrix, namely $\text{cond}(\mathbf{F}) \simeq 1$ everywhere.
3. The weak variation of $\|\mathbf{F}\|$ and $\text{cond}(\mathbf{F})$ on the celestial sphere, i.e. isotropic sensitivity to gws with uniform distribution of the polarization angle and isotropic reconstruction ability.
4. Similar sensitivity curves $S_{hh,1}(f) \dots S_{hh,M}(f)$, as very noisy detectors are effectively removed from the network and the maximum benefit is achieved when all detectors share approximately the same noise level in the same frequency band (e.g. it would be pointless to build a network of detectors with disjoint frequency bands).

We point out that the characterization provided by the “complete” figures of merit $\|\mathbf{F}\|$ and $\text{cond}(\mathbf{F})$ can be given as well by the angular-dependent power spectra of the $h_{\times}(t)$, $h_{+}(t)$ virtual channels defined by eq. 3.14. In fact, such power spectra play indeed the role of *network sensitivity curves*, i.e. the network extension of the single-detector $S_{hh}(f)$. It can be shown numerically that, in regions where $\text{cond}(\mathbf{F}) \simeq 1$, such sensitivity curves share a similar profile and even go below the single-detector ones. However, in regions where \mathbf{F} has a vanishing norm or it is strongly ill-conditioned, the network sensitivity curves tend to become very different and much larger than single-detector ones. This is another aspect of the reconstruction inability of a network with ill-conditioned response matrix.

In order to optimize the geographical factor of the response matrix, it must be noted that the gw propagation speed is better tested, and the source direction can be more accurately estimated, when the time delays $\tau_m(\vartheta, \varphi)$ differ greatly for every source direction, i.e. when the single-detector burst arrival times are widely spread around the mean arrival time. To achieve this, the detectors should have very different locations \mathbf{r}_m , with the obvious practical constraint that they must be located on the Earth surface. In fact, for a network of very close instruments, the delays vary strongly with (ϑ, φ) but they are always synchronized, i.e. the geographical factor of the response matrix becomes almost proportional to the identity matrix. If this is the case, the propagation speed is much more difficult to test and the estimation of (ϑ, φ) is more ambiguous, as it must rely only on the angular factor of \mathbf{F} , which has the coarse angular scale characteristic of single-detector antenna patterns. A useful figure of merit in this sense has been defined as the volume of the solid identified by the detector locations [72]. Accordingly, a good network should exhibit a fifth property.

5. The volume defined by the detector locations should be maximum. For instance, a 3-detector network defines a triangle whose area should be maximum; a 4-detector network identifies a prism whose volume should be maximum.

Clearly, a network that can hardly fulfill requirements 1-3 is composed of detectors with similar orientations. In fact, the zeroes of the single-detector antenna patterns all point in the same directions, creating a number of sky areas where $\|\mathbf{F}\| \simeq 0$, e.g. 2 areas for a network of bars and 4 for a network

of interferometers. The aligned antenna patterns also cause $\|\mathbf{F}\|$ to share their angular profile, thus the sensitivity is far from isotropic. Even worse, aligned antenna patterns compromise the conditioning, as the response matrix acquires nearly proportional rows almost everywhere in the sky. Row proportionality implies proportionality of the range vectors. It follows that the network response matrix is ill-conditioned almost everywhere. Interestingly, a network of aligned instruments is suitable for incoherent burst searches (e.g. IGEC), but it gives no solution to the gw inverse problem.

However, ill-conditioning may also arise if only two detectors of the network are close to alignment. One can show numerically that the condition number does not change if the quasi-aligned detector pair is replaced by a single instrument with sensitivity enhanced by a certain factor. Therefore, the network appears to effectively “lose” a detector and at the same time gain sensitivity in another one. This creates an unbalanced response matrix, with a detector significantly more sensitive than the others, which makes clear the origin of its ill-conditioning. The extreme case of *M exactly* aligned detectors is in fact equivalent to a single detector with higher sensitivity. The inability to reconstruct both gw amplitudes is then clear, as the full deconvolution can not be achieved with a single instrument.

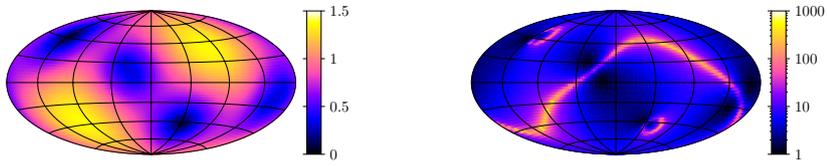
To study realistic cases, we can form a network with the most sensitive available interferometers, namely the two long-baseline LIGO detectors (Livingston and Hanford) and the VIRGO one. Numerical evaluation of the response matrix and network figures of merit² leads to the sky maps shown in figure 3.7. The top row presents a network with the two LIGO detectors only. Their strong alignment creates evident regions of the sky with small $\|\mathbf{F}\|$ and some with very large condition number, even larger than 10^3 . These problems are mitigated by the inclusion of VIRGO, which has a differently aligned antenna pattern. Adding the less sensitive GEO600 interferometer, which bears an even different alignment, heals the defects even more, notably the condition number whose maximum value drops by orders of magnitude. Note however that figure 3.7 only represents the *angular* factor of the network response matrix: as GEO600 has a poor sensitivity with respect to LIGO and VIRGO, the improvement is actually smaller than shown and acquires a frequency dependence.

Focusing on property 5, the upper plot of figure 3.3 tells us that the real network is not very satisfactory. In fact, although the current interferometers are fairly spread in longitude, they have been all built on the northern Earth hemisphere and with similar latitudes. As a result, they enclose a volume that is less than optimal for testing the gw propagation speed and for reducing the uncertainty in the estimated source location.

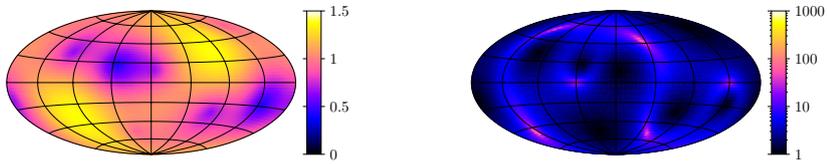
These examples lead to the intuition that a good network for studying gw bursts should use detectors with maximal misalignment *and* displacement. Indeed, we point out an optimal solution with respect to these requirements. This happens to be a particularly symmetric network, originally designed with the aim of estimating three gw invariants, namely the wave energy, the tensor trace and the determinant [73]. It is composed of six resonant bar

²For this purpose we developed tools running under the MATLAB environment.

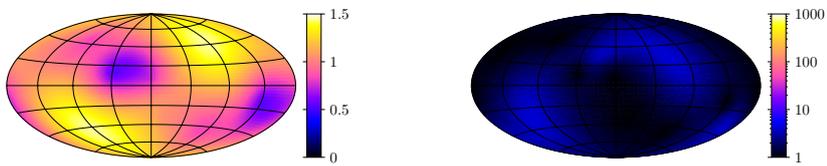
LIGO L, LIGO H



LIGO L, LIGO H, VIRGO



LIGO L, LIGO H, VIRGO, GEO600



Ideal network

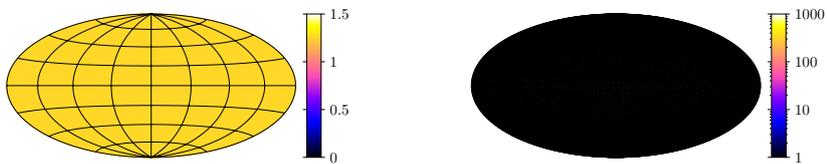


Figure 3.7: characteristics of the angular factor of the response matrix for several networks. The sky maps on the left show $\|\mathbf{F}(\vartheta, \varphi)\|$ while the right ones show the condition number $\text{cond}[\mathbf{F}(\vartheta, \varphi)]$. Note the improvement brought by the inclusion of VIRGO and GEO600 and the complete isotropy and perfect conditioning of the ideal network.

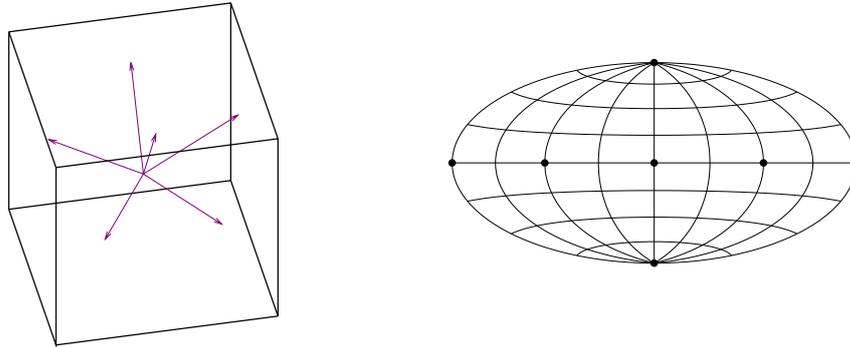


Figure 3.8: geometry of the ideal “dodecahedral” network. The left figure depicts the orientation of the six bar-like detectors. The vectors point from the origin to the centers of six faces of the regular dodecahedron. The right plot shows a desirable arrangement of the six detectors on the Earth surface.

detectors aligned along the axes of the regular dodecahedron (or, equivalently, pointing from the center to six vertices of the regular icosahedron) as sketched in figure 3.8. In the language of the network response matrix, we found that this setup remarkably exhibits *constant* figures of merit $\|\mathbf{F}\| = \sqrt{8/5}$ and $\text{cond}(\mathbf{F}) = 1$, as it can be seen in the bottom row of figure 3.7. The response matrix then provides isotropic sensitivity to gws and it can be perfectly inverted, allowing in principle the deconvolution of $h_+(t)$ and $h_\times(t)$ without problems. For these reasons, we will regard such a setup as the minimal “ideal” network with respect to properties 1-3. To satisfy property 4, the six detectors should simply have similar sensitivity curves. Finally, for property 5, their locations could be the six vertices of a regular octahedron inscribed in the Earth sphere.

Due to the known limits of bar detectors, clearly it would be more interesting an ideal network made of interferometers. We found several geometrical arrangements with very good figures of merit (e.g. $1 \leq \text{cond}(\mathbf{F}) < 2$ everywhere) although further research is needed for an interferometric equivalent of the “dodecahedral” bar network.

3.4.4 Geometrical interpretation of the network response

Eq. 3.8 provides a simple geometrical interpretation of a gw detection experiment performed with a network of detectors. Let the source direction (ϑ, φ) be fixed. $\mathbf{h}(t)$, an object with 2 degrees of freedom, is mapped into the M -dimensional space of $\mathbf{x}(t)$, which we call the *network output space*, by a simple linear application \mathbf{F} . This mapping takes place instant by instant and it is effectively independent from the time evolution of $\mathbf{h}(t)$. Assuming for simplicity that (ϑ, φ) is known, so that the time delays $\tau_m(\vartheta, \varphi)$ have been correctly compensated (or that the detectors are co-located, so the delays have in practice no effect) this means that the M -dimensional object

$\mathbf{F}\mathbf{h}(t)$ can only span a 2-dimensional *subspace* of the network output space. This happens regardless of the time dependence of $h_+(t)$ and $h_\times(t)$ or the source direction (ϑ, φ) . In other words, looking at the network output space, gws can only live on a *plane*, which we will call the *gw plane* for simplicity. The response of the network to gw signals follows a very strong geometrical constraint, particularly when $M \gg 1$ so that the network output space is much larger than the natural degrees of freedom of the gw polarization amplitudes.

The time evolution of $h_+(t)$ and $h_\times(t)$ merely fixes the trajectory traced by $\mathbf{h}(t)$ in its 2-dimensional space. This trajectory is then mapped by \mathbf{F} into the gw plane embedded in the network output space (figure 3.9). In the general case of complicated polarization (or no polarization at all) the samples of $\mathbf{h}(t)$ trace an arbitrary curve in its 2-dimensional space, and this curve is mapped into a planar figure in the network output space. Circularly polarized gws are characterized by polarization amplitudes in phase quadrature: this leads to a circular trajectory in the 2-dimensional space of $\mathbf{h}(t)$, which is mapped by \mathbf{F} into a planar trajectory (in general resembling an ellipse) constrained to the gw plane. A linearly polarized gw signal is characterized by the property $h_\times(t) \propto h_+(t)$: consequently, $\mathbf{h}(t)$ traces a line segment in its 2-dimensional space and $\mathbf{F}\mathbf{h}(t)$ must follow the same constraint, i.e. it traces a linear segment within the gw plane in the network output space. Different values of the polarization angle ψ correspond to different orientations of the line segment. Variations of ψ make these segments just span the whole gw plane.

In general, the orientation of the gw plane in the network output space depends on the components of \mathbf{F} , i.e. it is fixed by the antenna patterns and by the source direction. In fact, the 2-dimensional subspace associated with the gw plane is generated by the range vectors \mathbf{F}^+ , \mathbf{F}^\times . There is consequently a *set of planes* accessible to gws from arbitrary directions. Which planes are accessible depends on the antenna patterns and thus encodes some of the physical properties of gws. Waves sensed with different angular sensitivities, e.g. a possible scalar component of gws, would access a different portion of the output space. Clearly, due to signal sparsity, the samples of each gw burst populate only one of the planes at a time.

Physical signatures of gws also manifest themselves as transformations of the trajectory in the output space induced by transformations in the physical space. Notable examples are 90° and 180° rotations of the detector network about the gw wave vector. Due to the spin-2 character of gws, 90° rotations change the sign of the polarization states and 180° rotations leave the states unchanged (figure 1.1). Thus, the measured $h_+(t)$, $h_\times(t)$ waveforms must respectively change their sign and remain unchanged. It follows that the trajectory in the output space undergoes a reflection about the origin for 90° rotations and that it is invariant for 180° rotations.

Even the ill-conditioning of the response matrix has a simple interpretation within this geometrical discussion. \mathbf{F} becomes ill-conditioned when its range vectors get almost proportional. Accordingly, the gw plane is a subspace generated by quasi-aligned vectors. As such vectors represent nothing more than two basis vectors whose amplitudes are $h_+(t)$ and $h_\times(t)$, the result

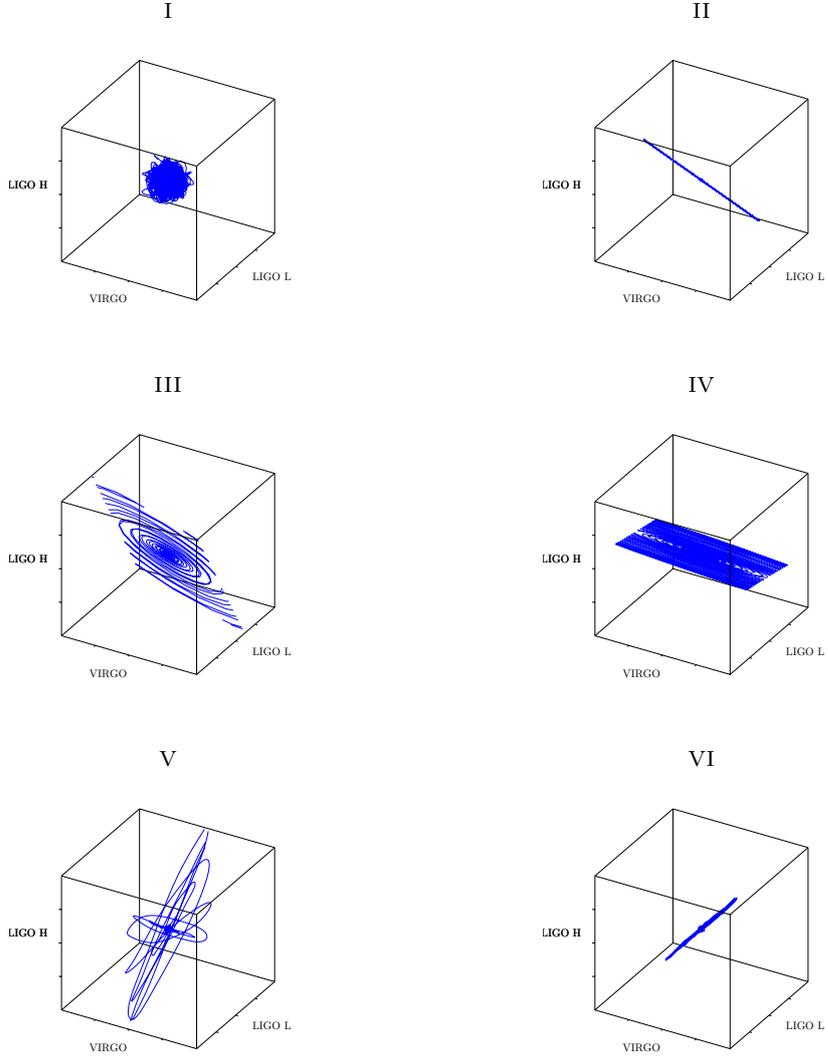


Figure 3.9: trajectories of simulated signals in the 3-dimensional LIGO-VIRGO network output space. The three axes represent the reconstructed data $x_m(t)$ delivered by each detector, correctly time-delayed to compensate the physical delays $\tau_m(\vartheta, \varphi)$. Time flows along the trajectory. I) Gaussian noise with equal variance in every detector. II) Linearly polarized sin-Gaussian gw burst. III) Circularly polarized sin-Gaussian gw burst. IV) The same burst coming from an ill-conditioned direction. V) Circularly polarized gw burst without $\tau_m(\vartheta, \varphi)$ correction. VI) Spurious glitch in LIGO L. The curves are clipped to the plot boxes to better show the planar structure of gws.

is that the shape of the trajectory traced by $\mathbf{h}(t)$, even if perfectly circular as in the case of circular polarization, is mapped into an extremely distorted, almost linear shape. In other words, the planar character of the gw tends to always degenerate into a *line*. To restore the planar structure, one of the two polarization amplitudes would have to be unphysically large. This phenomenon is nothing more than a manifestation of the rank deficiency of an ill-conditioned \mathbf{F} .

Unfortunately, for a network of distant detectors and unknown source direction, the de-synchronizing effect of the unknown time delays $\tau_m(\vartheta, \varphi)$ clearly distorts the planar shape of gws into a trajectory which, in general, spans an M -dimensional volume. This can be seen in the lower-left box of figure 3.9. To recognize the planar aspect of gws, then, one needs to scan the celestial sphere in search of a direction that flattens the trajectory to a plane.

The strong geometrical fingerprint exhibited by gws in correspondence of the correct time delay combination is to be compared with the behaviour of the spurious signals $\mathbf{s}(t)$ and the noise $\mathbf{n}(t)$. In the case of a single-detector glitch occurring in detector m , $\mathbf{s}(t)$ traces a line segment along the m axis of the network output space (figure 3.9). In the general case of glitches in more detectors occurring at the same time, $\mathbf{s}(t)$ may indentify a plane or a higher dimensional volume. As for the noise, it represents a stochastic process and consequently it spans the whole M -dimensional output space, with different magnitudes depending on each detector variance σ_m . In particular, if $\sigma_m = \sigma$ for $m = 1 \dots M$, the spanned volume is an M -sphere.

The discussed “degeneracy” of gw signals in the output space implies that, if $M > 2$, there exists a $(M - 2)$ -dimensional subspace, complementary to the gw plane, where there can be no gw contribution (*null space*) [60]. We may therefore construct a basis for the M -dimensional output space with 2 elements spanning the gw plane and the other elements spanning the null space [61]. For example, the LIGO-VIRGO network has a 1-dimensional null space, whose basis element can be constructed by the simple cross-product $\mathbf{F}_0 = \mathbf{F}^+ \times \mathbf{F}^\times$. Projecting $\mathbf{x}(t)$ along \mathbf{F}_0 then yields a time series where *any* gw signal coming from the correct (ϑ, φ) direction is removed, irrespectively of its time dependency. As anticipated, such a null time series is known as *null stream*.

If $M > 3$, more null streams are available and we need a more general way than the cross-product to construct them. This is provided as a “by-product” by the procedure leading to the Moore-Penrose inverse. In fact, one can define two $M \times M$ matrices

$$\begin{aligned} \mathbf{P} &= \mathbf{F}\mathbf{F}^\dagger \\ \mathbf{Q} &= \mathbf{I} - \mathbf{F}\mathbf{F}^\dagger \end{aligned} \tag{3.20}$$

where \mathbf{I} is the $M \times M$ identity matrix [68] (we omitted the dependencies of \mathbf{F} for simplicity, but \mathbf{P} and \mathbf{Q} clearly inherit them). The crucial property

of these matrices is that

$$\begin{aligned} \mathbf{P}\mathbf{F} &= \mathbf{F}\mathbf{F}^\dagger\mathbf{F} = \mathbf{F} \\ \mathbf{Q}\mathbf{F} &= \mathbf{F} - \mathbf{F}\mathbf{F}^\dagger\mathbf{F} = \mathbf{F} - \mathbf{F} = \mathbf{0} \end{aligned} \quad (3.21)$$

as $\mathbf{F}^\dagger\mathbf{F} = (\mathbf{F}^T\mathbf{F})^{-1}\mathbf{F}^T\mathbf{F} = \mathbf{I}$. Hence, \mathbf{P} leaves the network response to a gw signal unaltered, while \mathbf{Q} cancels it:

$$\begin{aligned} \mathbf{P}\mathbf{x}(t) &= \mathbf{P}(\mathbf{F}\mathbf{h}(t) + \mathbf{n}(t)) = \mathbf{F}\mathbf{h}(t) + \mathbf{P}\mathbf{n}(t) \\ \mathbf{Q}\mathbf{x}(t) &= \mathbf{Q}(\mathbf{F}\mathbf{h}(t) + \mathbf{n}(t)) = \mathbf{Q}\mathbf{n}(t). \end{aligned} \quad (3.22)$$

Similar straightforward algebra shows that $\mathbf{P}^2 = \mathbf{P}$, $\mathbf{Q}^2 = \mathbf{Q}$ and $\mathbf{Q}\mathbf{P} = \mathbf{0}$, and thus we recognize in \mathbf{P} and \mathbf{Q} a pair of *complementary and orthogonal projection matrices*. Their action on $\mathbf{x}(t)$ represents the projection on the gw plane and on the null space respectively. In other words, \mathbf{P} projects $\mathbf{x}(t)$ on the gw plane, leaving the gw signal unaltered and reducing the noise fluctuations, while \mathbf{Q} projects $\mathbf{x}(t)$ on the null space, removing the gw signal and also reducing the noise magnitude. Clearly, the sum of the two projections gives back $\mathbf{x}(t)$ as the projectors are complementary.

Other notable properties of \mathbf{P} and \mathbf{Q} arise straightforwardly from the fact that they are projectors. For the following discussion, relevant ones are their unit norm, namely $\|\mathbf{P}\| = \|\mathbf{Q}\| = 1$ and their eigenvalue spectra which reduce to $\{0, 1\}$. We stress that such properties follow from the projector character of the matrices in all generality and thus they can not depend on (ϑ, φ) .

The unit norm notably implies that no divergency can occur when applying \mathbf{P} and \mathbf{Q} to the data, even when \mathbf{F} is ill-conditioned. This is because a projection means asking the inverse problem much less than the full reconstruction of $h_+(t)$ and $h_\times(t)$.

Let us summarize the discussion. Naturally, gw signals contribute to the detector output only along a 2-dimensional plane in the network output space, whose orientation depends on the source direction. For an arbitrary network configuration with $M > 2$, the procedure for constructing the Moore-Penrose inverse provides a way to project the observed data on the gw plane and on the complementary null space. This defines two new synthetic time series $\mathbf{x}_P(t) = \mathbf{P}\mathbf{x}(t)$ and $\mathbf{x}_Q(t) = \mathbf{Q}\mathbf{x}(t)$ out of $\mathbf{x}(t)$: the former carries reduced noise but identical gw signal, the latter carries reduced noise only. Differently from what can happen with the reconstructed $h_+(t)$ and $h_\times(t)$ synthetic streams, $\mathbf{x}_P(t)$ and $\mathbf{x}_Q(t)$ are always characterized by a finite noise variance, which makes them more suitable for tackling the detection and discrimination problems.

3.5 Detection and discrimination through geometric projections

Thanks to the definition of $\mathbf{x}_P(t)$ and $\mathbf{x}_Q(t)$, the geometrical constraint provides in principle a way to distinguish gw bursts from noise and glitches. In fact, we can scan the celestial sphere and for each direction compute such

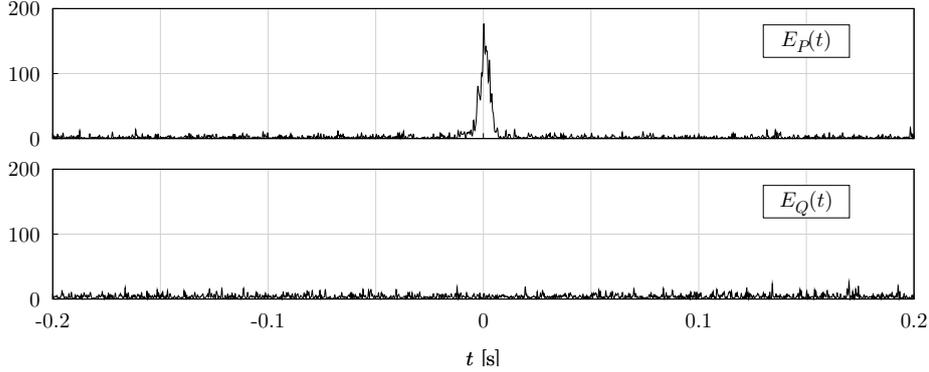


Figure 3.10: time evolution of the E_P and E_Q channels in presence of a gw burst, using the ideal network. E_P and E_Q are evaluated at the correct source direction. The gw burst is identified as the excess of energy in the first plot with respect to the second.

time series, looking for a direction for which $\mathbf{x}_Q(t)$ loses any possible gw signal with respect to $\mathbf{x}_P(t)$ [61]. This should be the signature of a planar structure of the data, useful for recognizing gws among noise and glitches. As a by-product, the scan also gives us an estimate of the source direction.

The M components of $\mathbf{x}_P(t)$ and $\mathbf{x}_Q(t)$ are no longer associated with a particular detector, as they represent linear combinations of the single detector outputs. There is no reason to treat each component differently from the others, thus we gather such M -dimensional information into two more convenient statistics, the quadratic forms

$$\begin{aligned} E_P(t) &= \|\mathbf{x}_P(t)\|^2 = \mathbf{x}^T(t) \mathbf{P} \mathbf{x}(t) \\ E_Q(t) &= \|\mathbf{x}_Q(t)\|^2 = \mathbf{x}^T(t) \mathbf{Q} \mathbf{x}(t) \end{aligned} \quad (3.23)$$

where we exploited the property $\mathbf{P}^T \mathbf{P} = \mathbf{P}$ and $\mathbf{Q}^T \mathbf{Q} = \mathbf{Q}$. E_P and E_Q represent the instantaneous “energy” of the P and Q projections (the “gw energy” and the “null energy”) or more precisely the magnitude of the data vector projected in the two complementary subspaces. Note that the E_P , E_Q time series depend on the direction (ϑ, φ) we are considering. In particular, E_Q is the *residual function* of the inverse problem and it can be used to find the true gw source direction [68] as will be clear in the following. As an example, figure 3.10 shows the E_P , E_Q time series for a gw burst, evaluated at the true source direction.

To identify the planar structure of a possible gw burst we can then evaluate E_P and E_Q over the celestial sphere and look for time intervals and directions where there is an excess of energy in the former, but significantly less energy in the latter. In fact, this is the signature of a vector belonging to one of the gw planes. Figure 3.11 shows the sky maps of one sample of E_P and E_Q , for a data set containing only Gaussian noise and for another data set also containing a gw burst. The oscillations of the gw burst waveforms produce interference fringes in the sky maps, because, due to the time delays $\tau_m(\vartheta, \varphi)$, moving on the celestial sphere effectively shifts the M realizations

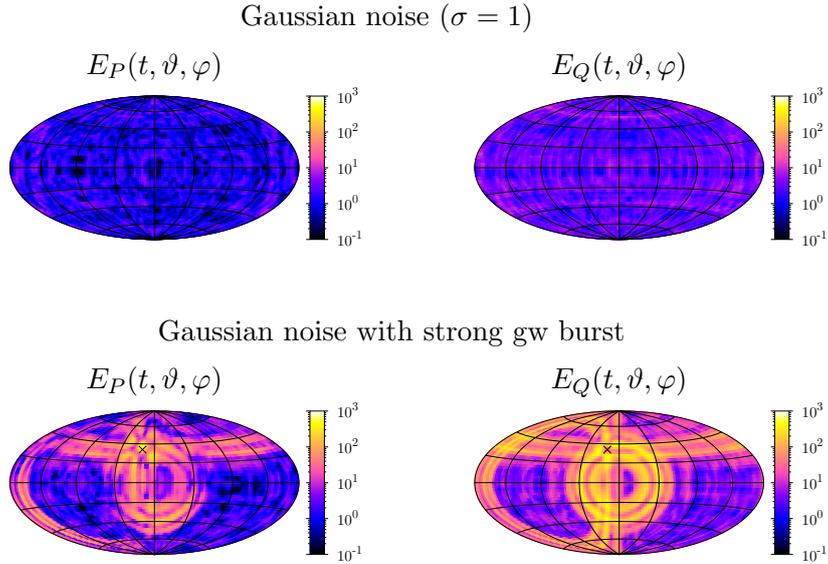


Figure 3.11: sky maps of one sample of the virtual channels E_P , E_Q ($t = 0$) for the ideal network. In the bottom sky maps, the source direction is marked by the ‘ \times ’. Note the (mean) isotropy of the noise in the top sky maps, and the interference fringes due to the time delays in the bottom ones.

of the signal alternatively in and out of phase. In particular, the sky map is populated with M “rings”, each corresponding to one of the detectors, that produce constructive and destructive interference. The rings fly over the celestial sphere as time passes and they all intersect in a single spot when the mean wavefront of the burst crosses the center of the Earth. The spot matches the true source direction (the ‘ \times ’ in figure 3.11) and in its neighborhood we find an excess of energy in E_P associated with a reduction of energy in E_Q .

3.5.1 Statistics of E_P and E_Q

E_P and E_Q are described by remarkably simple statistics. In fact, under reasonable assumptions about the noise, the properties of \mathbf{P} and \mathbf{Q} allow us to derive the joint probability density $P(E_P, E_Q)$.

Let us start by writing

$$P(E_P, E_Q) = \int P(\mathbf{x}) \delta(E_P - \mathbf{x}^T \mathbf{P} \mathbf{x}) \delta(E_Q - \mathbf{x}^T \mathbf{Q} \mathbf{x}) d\mathbf{x} \quad (3.24)$$

where $P(\mathbf{x})$ is the probability density of one sample of the network data vector \mathbf{x} . The two Dirac deltas can be written using their Fourier transforms,

$$P(E_P, E_Q) = \iiint P(\mathbf{x}) \exp [uE_P + vE_Q - \mathbf{x}^T (u\mathbf{P} + v\mathbf{Q}) \mathbf{x}] du dv d\mathbf{x} \quad (3.25)$$

where the integrals in du and dv are performed along the imaginary axis. Now assume the noise is Gaussian distributed, uncorrelated between differ-

ent detectors and with identical variance σ^2 in every detector, namely

$$P(\mathbf{x}) = \frac{1}{(2\pi\sigma^2)^{\frac{M}{2}}} \exp\left(-\frac{1}{2\sigma^2}(\mathbf{x} - \mathbf{s})^T(\mathbf{x} - \mathbf{s})\right) \quad (3.26)$$

where \mathbf{s} is any signal (either gw or spurious) superposed to the noise. Then,

$$\begin{aligned} P(E_P, E_Q) &= \left(\frac{\alpha}{\pi}\right)^{\frac{M}{2}} \iiint \exp[-\alpha(\mathbf{x} - \mathbf{s})^T(\mathbf{x} - \mathbf{s})] \cdot \\ &\quad \cdot \exp[-\mathbf{x}^T(u\mathbf{P} + v\mathbf{Q})\mathbf{x}] d\mathbf{x} \cdot \\ &\quad \cdot e^{uE_P} e^{vE_Q} du dv, \end{aligned} \quad (3.27)$$

where $\alpha = 1/2\sigma^2$. Writing \mathbf{x} as $\mathbf{s} + \mathbf{n}$ and switching the integration variable to \mathbf{n} yields

$$\begin{aligned} P(E_P, E_Q) &= \left(\frac{\alpha}{\pi}\right)^{\frac{M}{2}} \cdot \\ &\quad \cdot \iiint \exp[-\mathbf{n}^T(\alpha\mathbf{I} + u\mathbf{P} + v\mathbf{Q})\mathbf{n} - 2\mathbf{n}^T(u\mathbf{P} + v\mathbf{Q})\mathbf{s}] d\mathbf{n} \cdot \\ &\quad \cdot \exp[-\mathbf{s}^T(u\mathbf{P} + v\mathbf{Q})\mathbf{s}] e^{uE_P} e^{vE_Q} du dv. \end{aligned} \quad (3.28)$$

The integration in $d\mathbf{n}$ can be done by noting that it is a standard M -dimensional Gaussian integral with the linear term, and in general, for any $M \times M$ symmetric matrix \mathbf{A} and M -vector \mathbf{b} ,

$$\int \exp(-\mathbf{n}^T \mathbf{A} \mathbf{n} + \mathbf{b}^T \mathbf{n}) d\mathbf{n} = \frac{\pi^{M/2}}{\sqrt{\det(\mathbf{A})}} \exp\left(\frac{\mathbf{b}^T \mathbf{A}^{-1} \mathbf{b}}{4}\right). \quad (3.29)$$

In our case,

$$\begin{aligned} \mathbf{A} &= \alpha\mathbf{I} + u\mathbf{P} + v\mathbf{Q} \\ \mathbf{b} &= 2(u\mathbf{P} + v\mathbf{Q})\mathbf{s}. \end{aligned} \quad (3.30)$$

Now we exploit the properties of \mathbf{P} and \mathbf{Q} . Using their complementarity, we can write

$$\mathbf{A} = (\alpha + u)\mathbf{P} + (\alpha + v)\mathbf{Q} \quad (3.31)$$

and from the fact that they are orthogonal and idempotent we also have

$$\mathbf{A}^{-1} = (\alpha + u)^{-1}\mathbf{P} + (\alpha + v)^{-1}\mathbf{Q}, \quad (3.32)$$

hence

$$\mathbf{b}^T \mathbf{A}^{-1} \mathbf{b} = 4 \left(\frac{u^2}{\alpha + u} \mathbf{s}^T \mathbf{P} \mathbf{s} + \frac{v^2}{\alpha + v} \mathbf{s}^T \mathbf{Q} \mathbf{s} \right). \quad (3.33)$$

Furthermore, as \mathbf{P} and \mathbf{Q} are projection matrices, their eigenvalues are $\{0, 1\}$ with multiplicities respectively $\{M - 2, 2\}$ for \mathbf{P} and $\{2, M - 2\}$ for \mathbf{Q} . Then, writing \mathbf{A} in diagonal form is trivial and leads to

$$\det(\mathbf{A}) = (\alpha + u)^2(\alpha + v)^{M-2}, \quad (3.34)$$

determinants being independent from the basis. By using 3.33 and 3.34 in 3.29 one can see that the Gaussian integral splits into the product of

factors involving either u or v . By further substituting in 3.28, the remaining integrals separate and the probability density remarkably factorizes as

$$P(E_P, E_Q) = P(E_P) P(E_Q) \quad (3.35)$$

with

$$\begin{aligned} P(E_P) &= \int \frac{1}{1 + 2\sigma^2 u} \exp\left(\frac{-s_p u}{1 + 2\sigma^2 u}\right) e^{u E_P} du \\ P(E_Q) &= \int \left(\frac{1}{1 + 2\sigma^2 v}\right)^{\frac{M}{2}-1} \exp\left(\frac{-s_q v}{1 + 2\sigma^2 v}\right) e^{v E_Q} dv \end{aligned} \quad (3.36)$$

and $s_p = \mathbf{s}^T \mathbf{P} \mathbf{s}$, $s_q = \mathbf{s}^T \mathbf{Q} \mathbf{s}$. As a final step, recall that these integrals represent two inverse Fourier transforms. The transformed functions now match the moment generating functions of two non-central χ^2 distributions, with 2 and $M - 2$ degrees of freedom respectively, and whose non-centrality parameters are $\mathbf{s}^T \mathbf{P} \mathbf{s}$ and $\mathbf{s}^T \mathbf{Q} \mathbf{s}$ respectively. Thus,

$$\begin{aligned} P(E_P) &= \frac{1}{2} \exp\left(-\frac{E_P + s_p}{2\sigma^2}\right) I_0\left(\frac{\sqrt{E_P s_p}}{\sigma^2}\right) \\ P(E_Q) &= \frac{1}{2} \exp\left(-\frac{E_Q + s_q}{2\sigma^2}\right) \left(\frac{E_Q}{s_q}\right)^{\frac{M}{4}-1} I_{M/2-2}\left(\frac{\sqrt{E_Q s_q}}{\sigma^2}\right) \end{aligned} \quad (3.37)$$

where $I_\nu(a)$ is the modified Bessel function of the first kind.

Summarizing, under the assumption of white Gaussian noise, uncorrelated between different detectors and with identical variance, several notable conclusions can be drawn about the virtual channels E_P and E_Q , which make them interesting for the detection and discrimination of gws.

1. E_P is distributed as a non-central χ^2 with 2 degrees of freedom and non-centrality parameter equal to $\mathbf{s}^T \mathbf{P} \mathbf{s}$, i.e. the magnitude of the signal projection in the P subspace.
2. E_Q is distributed as a non-central χ^2 with $M - 2$ degrees of freedom and non-centrality parameter equal to $\mathbf{s}^T \mathbf{Q} \mathbf{s}$, i.e. the magnitude of the signal projection in the Q subspace.
3. E_P and E_Q are *statistically independent* processes.

In absence of any signal, i.e. $\mathbf{s} = 0$, the virtual channels exhibit a notably simple behaviour. In fact, the two distributions reduce to independent central χ^2 , with 2 and $M - 2$ degrees of freedom (figure 3.12). Then, as \mathbf{P} and \mathbf{Q} only appear in the non-centrality terms, the virtual channels become—on average— independent both from the angles (ϑ, φ) and the network geometry. We can see that in the top sky maps of figure 3.11: $P(E_P, E_Q)$ remains unchanged while we scan the celestial sphere. Such invariance can be interpreted geometrically by recalling that the network noise, assuming the same variance in every detector, identifies a spherical volume in the network output space. Then, slicing this volume with differently oriented

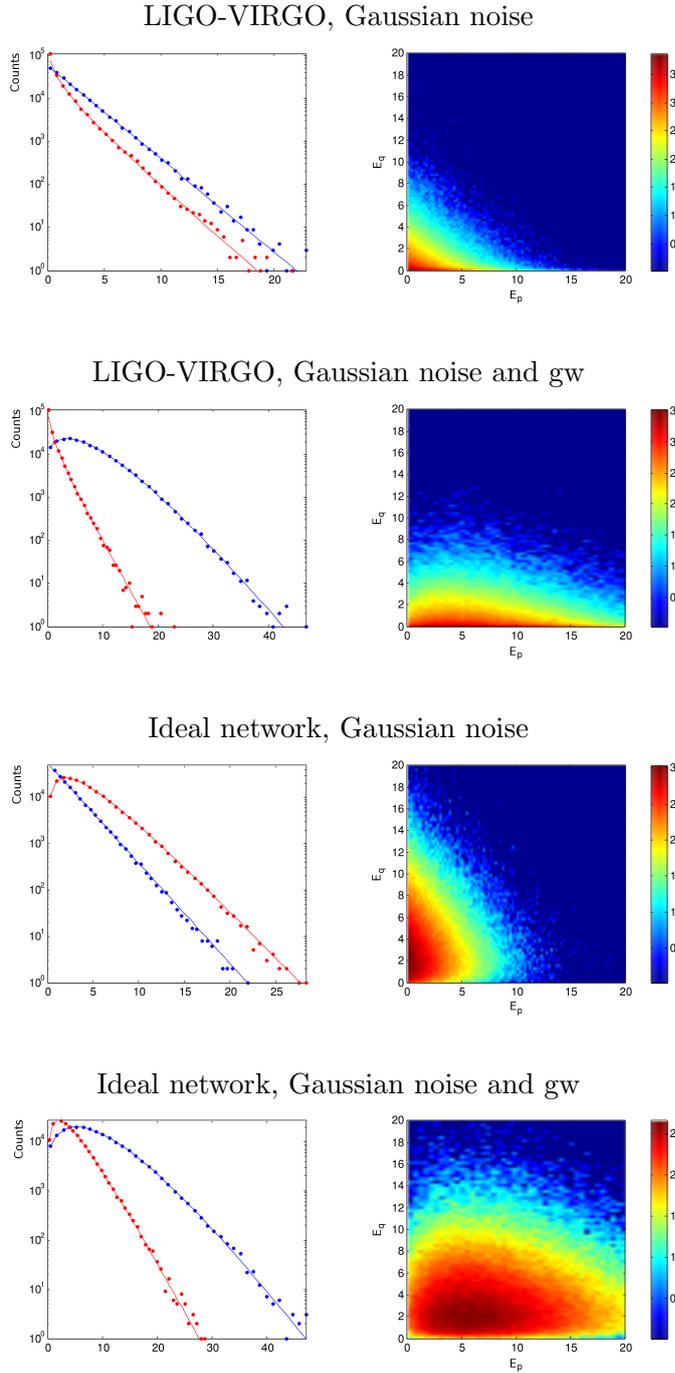


Figure 3.12: results of simple Monte Carlo simulations to evaluate the statistics of one sample of E_P and E_Q , in presence of noise only and noise+gw. The left plots are the histograms of E_P (blue dots) and E_Q (red dots) while the right plots sketch the 2-dimensional histograms of (E_P, E_Q) (log color scale). For the gw cases, E_P and E_Q are evaluated at the correct source direction. The histograms follow the theoretical $P(E_P)$ and $P(E_Q)$ (continuous lines in the left plots).

planes (i.e. projecting it with \mathbf{P} and \mathbf{Q} for different source directions) always yields statistically the same result. Clearly, a dependence on (ϑ, φ) appears as soon as the noise variances in the detectors are different, because the noise trajectory in the output space loses its mean isotropy and becomes an M -dimensional ellipsoid.

The possible ill-conditioning of \mathbf{F} does not have an explicit effect in these statistics. In fact, it only causes the collapse of the 2-dimensional gw trajectory towards a 1-dimensional segment and no divergence can occur if this curve is projected with \mathbf{P} and \mathbf{Q} . We only expect a reduction of energy in E_P with respect to the gw energy h_{rss} and accordingly a difficulty to detect and recognize gw bursts coming from ill-conditioned directions.

When \mathbf{s} represents a non-null gw contribution and the virtual channels are evaluated at the correct source direction, \mathbf{Q} cancels the non-centrality of E_Q , $P(E_Q)$ reduces to a central χ^2 and E_Q attains, on average, its minimum. On the other hand, the signal energy is forced to concentrate in the P channel and $P(E_P)$ reaches its maximum non-centrality. This accounts for the significant suppression of E_Q with respect to E_P (figure 3.12).

When \mathbf{s} is a gw signal, but we are evaluating the virtual channels at wrong source angles, there is no cancellation in E_Q . This also happens when \mathbf{s} is a spurious signal. In general, gws from wrong directions and spurious signals can not be directly distinguished in the virtual channels. The difference lies in the fact that, for gws, one can always find a direction in the celestial sphere where a strong E_Q cancellation occurs with respect to E_P , while for spurious signals such a direction does not, in general, exist.

Nevertheless, a network could be “degenerate” in the sense that some of the gw planes may contain one of the axes. This occurs whenever a direction exists for which the antenna patterns of all detectors vanish except one. Such degeneracy is troublesome because, if this is the case, gws are truly indistinguishable from single spurious signals produced by the corresponding detector, whatever the analysis algorithm is. In fact, *any* transient signal in that detector can always be interpreted as a gw burst coming from the problematic direction. This serious issue does indeed happen for the LIGO-VIRGO network: for two opposite sky patches, two detectors have almost null response to gws while the third is more sensitive (figure 3.13). However, these patches are fairly small and they can always be excluded from the gw

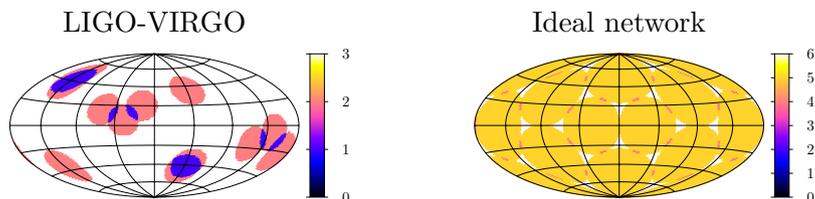


Figure 3.13: sky maps of the number of detectors for which $\hat{F}(\vartheta, \varphi) > 0.2$. In the LIGO-VIRGO network, note the patches where only a single detector is sensitive (blue areas).

search, i.e. the visible sky can be slightly reduced to “safe” regions only. Equivalently, one may always veto any gw burst detected by the analysis and coming from those directions. Figure 3.13 also shows that the ideal dodecahedral network does not exhibit such pathologies and covers the whole sky with a uniform number of “active” instruments.

3.5.2 Time-frequency multiresolution analysis

In the above discussion, the samples of E_P and E_Q are treated independently of each other, as \mathbf{x} can be effectively projected instant by instant. Consequently, searching the sky for an excess of energy in E_P matching a lack of energy in E_Q can be performed instant by instant. Clearly, we are not taking into account the expected phase coherence of the gw burst. Accordingly, by using the time domain only we are spreading the unknown time-frequency structure of the gw waveform over a possibly large number of samples. On the contrary, we would like to exploit the expected waveform phase coherence to better recognize the gw burst buried in the noise. In particular, it would be useful to concentrate the signal in a limited number of high energy samples (ideally just one), as this would make the difference between E_P and E_Q stand higher above the noise fluctuations.

For this purpose we could use intermediate time-frequency representations of \mathbf{x} rather than the time-domain only. A suitable tool for the time-frequency representation of unknown transient signals is the *wavelet transform* and in particular its discrete implementation known as the *wavelet packet decomposition* [74]. In the following, we give a brief introduction to such formalism.

The *continuous wavelet transform* decomposes a signal on basis functions constructed by translating and scaling a unique prototype function $\psi(t)$ (*mother wavelet*). In fact, translation and scaling operations can be shown to provide a basis for e.g. the $L^2(\mathbb{R})$ space of functions, where we assume to find the gw waveforms. The basis functions are waveforms of limited duration (the opposite of the sinusoidal basis of the Fourier analysis) and limited bandwidth (the opposite of the δ basis of the time domain). Such feature realizes the simultaneous time and frequency localization of the signal spectral components (to an extent compatible with the Heisenberg uncertainty principle, clearly). This leads to the key feature of the wavelet analysis, i.e. its efficient representations of transient signals. The basis functions have the form

$$\psi_{\tau,s}(t) = \frac{1}{\sqrt{|s|}} \psi\left(\frac{t-\tau}{s}\right) \quad (3.38)$$

with $\tau, s \in \mathbb{R}$. τ and $s \neq 0$ represent the location and scale parameters, which control the time-frequency extension of the basis element. By varying τ and s one can partition the full time-frequency plane with different resolutions. The continuous wavelet transform of a signal $x(t)$ takes the form of a simple scalar product in $L^2(\mathbb{R})$,

$$\mathcal{X}(\tau, s) = \int_{\mathbb{R}} \psi_{\tau,s}^*(t) x(t) dt. \quad (3.39)$$

It can be also shown that under a suitable *admissibility condition* about $\psi(t)$, $x(t)$ can be fully reconstructed from $\mathcal{X}(\tau, s)$. Many classes of mother wavelets $\psi(t)$ have been introduced, each with characteristic properties. There is no strict indication for the choice of the mother wavelet and one is generally guided by the characteristics of the signals to analyze, e.g. continuity and differentiability for gw bursts. In our simulations we use high order wavelets of the *symlet* family, i.e. the symlet 10.

The continuous wavelet transform is strongly redundant and it can be discretized without loss of information. Sampling τ and s on a dyadic mesh (i.e. $\tau = j2^{-k}$ and $s = 2^{-k}$ with $j, k \in \mathbb{Z}$) yields the *discrete wavelet transform* (DWT) [74]. A notable feature of the DWT is its connection with the *multiresolution analysis* (MRA) [74]. The MRA is the decomposition of a signal into a hierarchy of functions that approximate it more and more accurately. Each function lies within a closed *approximation subspace* $V_k \subset L^2(\mathbb{R})$ so that $V_{k-1} \subset V_k \subset V_{k+1}$ and

$$\lim_{k \rightarrow -\infty} V_k = \{0\}, \quad \lim_{k \rightarrow +\infty} V_k = L^2(\mathbb{R}). \quad (3.40)$$

In other words, the index k determines the resolution of the approximation, i.e. the function in V_{k+1} better approximates the signal than the one in V_k . In addition, the MRA defines the *detail subspace* W_k as the orthogonal complement of V_k in V_{k+1} . W_k contains the additional detail required to improve the resolution from level k to $k+1$. An orthogonal basis for each V_k is provided by the discrete translations of a *scaling function* $\phi(t)$, i.e. $\{\phi_{k,j}\} = \{2^{k/2}\phi(2^{k/2}t - j)\}$. Conversely, an orthogonal basis for W_k is provided by translations of a wavelet $\psi(t)$, namely $\{\psi_{k,j}\} = \{2^{k/2}\psi(2^{k/2}t - j)\}$. The coefficients of such bases are the *approximation* and *detail* coefficients respectively. Note that, due to the complementarity of the approximation and detail subspaces, both $\{\phi_{k+1,j}\}$ and $\{\phi_{k,j}, \psi_{k,j}\}$ are bases for V_{k+1} .

The connection between DWT and MRA arises because the sampling process represents a continuous signal $x(t)$ by discrete samples $\{x[j]\} = \{x(jt_s)\}$, where t_s is the sampling period. If t_s is small enough, we can write $t_s = 2^{-k}$ with k a large integer. Therefore, the sample set $\{x[j]\}$ can be thought as the MRA approximation coefficients of $x(t)$ within some approximation space V_0 . Thus, we can exploit the hierarchy of approximation and detail subspaces and decompose

$$V_0 = V_{k_0} \oplus W_{k_0} \oplus W_{k_0+1} \oplus \dots \oplus W_{-1}. \quad (3.41)$$

Accordingly, the scaling functions for level $k = k_0$ and the wavelets for levels $k = k_0 \dots -1$ form together an orthogonal basis for V_0 . In such a framework, the DWT of $x(t)$ is simply the expansion over this basis, i.e. the set of approximation coefficients $\{a_{k_0,j}\}$ and detail coefficients $\{\{d_{k_0,j}\} \dots \{d_{-1,j}\}\}$.

The crucial feature that makes the DWT practically feasible is the fact that one does not need to explicitly calculate the scaling functions, the wavelets and their scalar products with the signal. In fact, $\phi(t)$ and $\psi(t)$ can be defined through a pair of discrete finite impulse response (FIR) filters, called *conjugate mirror filters*, that represent two complementary band-pass

operations. The evaluation of the DWT is then equivalent to the iterative application of such filters to the data samples, followed by a decimation of the filtered samples at each iteration. This is known as the Mallat algorithm and it is computationally very efficient, playing the role of the fast Fourier transform algorithm for the discrete Fourier transform.

Since each basis element can be associated with a particular (finite) time-frequency support, the whole basis of the DWT determines a characteristic partition or “tessellation” of the time-frequency plane, as it is also true for other bases like the time domain and Fourier basis. The time domain is associated with tiles that are infinitesimally thin along the time axis and infinitely extended along the frequency one: this accounts for the well-known absence of frequency resolution of the time domain. On the opposite extreme is the Fourier domain, which partitions the time-frequency plane into tiles infinitesimally thin along the frequency axis and infinitely extended along the temporal dimension. Instead, the DWT partition is made of constant-area tiles with better frequency resolution at low frequency and better time resolution at high frequency. Such transition in resolution resembles the logarithmic behaviour of the dyadic grid with which τ and s are sampled.

Although the partition associated with the DWT is very useful for slow signals contaminated by fast fluctuations in the noise, actually there is no reason for such variation of the time-frequency resolution in gw burst searches: in fact—as discussed in chapter 2—gw waveforms pass through the detectors, which limit their time-frequency extension to finite frequency bands. Given the unknown time-frequency structure of gw bursts, a uniform partition of the time-frequency plane is therefore preferred.

This is provided by a computationally efficient generalization of the DWT, known as *wavelet packet* (WP) decomposition [74]. The WP decomposition generalizes the DWT by not only decomposing each approximation subspace V_k into V_{k-1} and W_{k-1} , but also by decomposing the detail subspace W_k . This decomposition can be arbitrarily iterated and continued, deciding at each iteration which subspaces to further split (figure 3.14). This allows one to achieve arbitrary partitions of the time-frequency plane. For

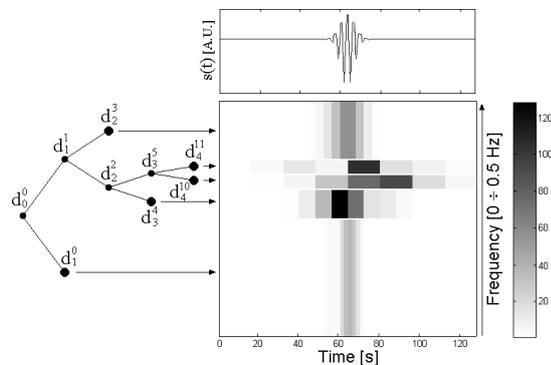


Figure 3.14: example wavelet packet decomposition. On the left, the associated binary tree is shown.

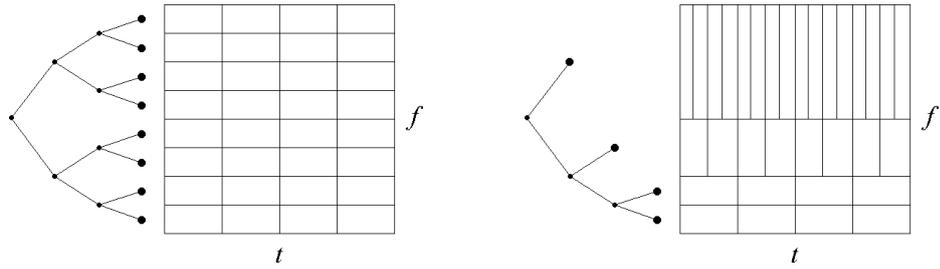


Figure 3.15: on the left, partition of the time-frequency plane associated with a complete wavelet packet decomposition. On the right, the partition associated with the discrete wavelet transform.

example, one may decompose every subspace at each iteration, obtaining an uniform partition of the time-frequency plane (complete decomposition). Or, one may just decompose the V_k subspaces, recovering a DWT-like partition. Each partition of the time-frequency plane is uniquely associated with a chain of iterative decompositions and it can be conveniently represented by a binary tree, whose nodes represent the subspaces (figure 3.15). The leaves are subspaces that are no furtherly decomposed. They are mutually orthogonal and their direct sum gives back the original space (the root of the tree).

The complete WP decomposition is suitable to explore the available data at more time-frequency resolutions. In fact, each successive decomposition doubles the frequency resolution and halves the time one, further departing from the time domain and approaching the Fourier basis more and more. It has been shown by Monte Carlo simulations that a threshold on the maximum coefficient of a complete WP decomposition defines a decision rule with very good detection efficiency for single-detector gw burst searches [54].

Thus, we can exploit the multiresolution analysis provided by the WP transform by decomposing the M data sets $\{\mathbf{x}[j]\}$ up to a maximum level K , producing the vector of coefficients $\{\mathbf{d}_k^p[j]\}$, where p labels the time-frequency pixel along the frequency axis, j labels it along the time axis and k labels the decomposition level up to K . Including the time domain, this gives rise to K redundant representations of $\{\mathbf{x}[j]\}$ that scan the time-frequency plane with different resolutions, starting with the maximum resolution in time (the time domain itself) and ending with a better resolution in frequency, as we approach the maximum achievable decomposition level (fixed by the limited number of available samples N). Among these different partitions, there is one whose tiles best match the time-frequency extension of the gw signal. Then, as the different decomposition levels can be shown to be orthogonal, the signal concentrates into few high energy coefficients, while the noise is spread among the other coefficients.

By constructing the E_P , E_Q statistics from the vector of WP coefficients $\{\mathbf{d}_k^p[j]\}$, we thus create K time-frequency maps of gw and null energies, functions of (ϑ, φ) . Thanks to the linearity of the WP decomposition, Gaussian noise in the data implies that each coefficient is also Gaussian distributed

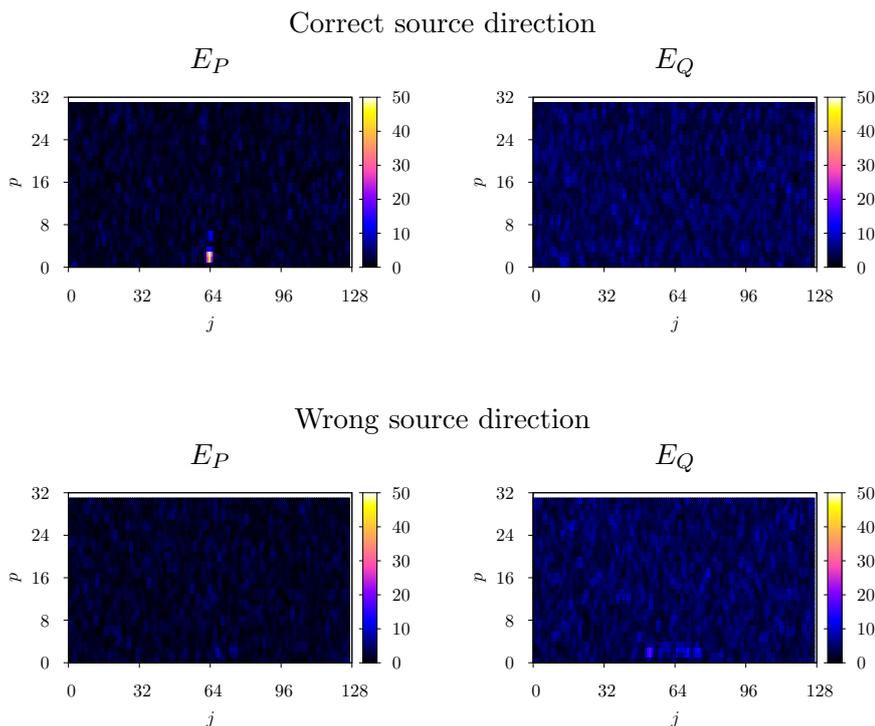


Figure 3.16: wavelet packet time-frequency maps of E_P and E_Q , at the decomposition level $k = 6$, for a simulated gw burst seen by the ideal detector network. In the top plots, note the complete vanishing of the gw time-frequency structure on the E_Q map. Such disappearance can be hardly noticed when the source direction is wrong.

and thus E_P , E_Q retain the χ^2 distributions discussed previously. These “extended” E_P , E_Q maps allow us to achieve better localization and recognition of any excess of energy in E_P with respect to E_Q . As an example, we sketch in figure 3.16 the $k = 6$ WP maps of E_P and E_Q . The simulated gw burst, though of modest amplitude with respect to the noise standard deviation, is strongly concentrated in a few intense pixels, allowing one to readily recognize its cancellation in E_Q . Failing to match the correct source direction evidently disrupts the synchronization of the time delays, spreading the burst over several weaker coefficients. Furthermore, the mismatched projections tend to distribute the weak residual energy equally in both maps.

It is worth noticing that, if the time-frequency support and source location of the gw burst were known, we could avoid scanning the time-frequency plane at different resolutions and searching the celestial sphere. In fact, we would be able to localize the burst energy, both in time-frequency and source direction. This would be equivalent to an “extended” local power test and thus a threshold on E_P would provide an *optimal* detection method with respect to the Neyman-Pearson criterion. At the same time, a null detection on E_Q would provide a strong discrimination of non-gravitational signals. Lacking such prior information, we must resort to scanning the complicated

map on the time-frequency plane and celestial sphere, looking for an excess of energy in E_P with respect to E_Q .

3.5.3 The PQ plane

So far we pointed out how we can recognize the planar structure of gw signals by gathering the original M -dimensional information into two synthetic, power-like statistics E_P and E_Q . At this point, detection and discrimination rules must be defined on such statistics as required by the initial discussions and in particular by figure 3.2.

We formulate this problem by considering the 2-dimensional (E_P, E_Q) plane (PQ plane in short). As the time-frequency plane and celestial sphere are scanned, the PQ plane is populated differently by the collected data. In particular, we can split it into three regions accordingly to the joint probability density $P(E_P, E_Q)$.

One region is mainly populated when the data contain Gaussian noise only. In this case, $P(E_P, E_Q)$ is the product of two central χ^2 distributions with 2 and $M - 2$ degrees of freedom. As M is always a small number, $P(E_P, E_Q)$ concentrates mainly near the origin of the PQ plane, with exponentially-decaying tails towards larger E_P and E_Q (figure 3.17, box I). The independence of $P(E_P, E_Q)$ from (ϑ, φ) implies that the Gaussian noise is constrained into this region while we scan the celestial sphere, provided that the noise variance has been correctly normalized between the detectors.

The second region is populated by gw signals that come from the direction (ϑ, φ) at which E_P and E_Q are evaluated. In fact, in this case $P(E_P)$ becomes non-central and its bulk shifts towards larger E_P values, while $P(E_Q)$ is still a central χ^2 . Thus, gws from the tested direction populate a horizontal “strip” bounded from below by the E_P axis (figure 3.17, box I).

The rest of the PQ plane is populated when the non-centralities of both statistics become significant, i.e. for spurious signals and gws coming from a wrong sky direction. This stresses even more the similar role played by these signal classes.

In presence of a signal, the non-centralities of $P(E_P, E_Q)$ keep the bulk of the distribution away from the noise region. Nevertheless, the non-centralities become differently distributed among the P and Q subspaces while we scan the sky (we are slicing the network output space with differently oriented planes). Accordingly, the data oscillates between the E_P and E_Q axes. How close it gets to the gw region can be used to discriminate if the data represents a gw burst or a spurious signal, because spurious signals will never get into the gw region. In other words, gw detection and discrimination are performed simultaneously by characterizing how the distribution of the data moves on the PQ plane while the time-frequency plane and celestial sphere are explored.

Therefore, as the final step to construct our detection and discrimination algorithm, we need to define the region of gw acceptance in the PQ plane. According to the statistics of E_P and E_Q , a simple solution is the rectangular region bounded by two thresholds T_1 and T_2 , i.e.

$$R_1 = \{(E_P, E_Q) \mid E_P > T_1, E_Q < T_2\} \quad (3.42)$$

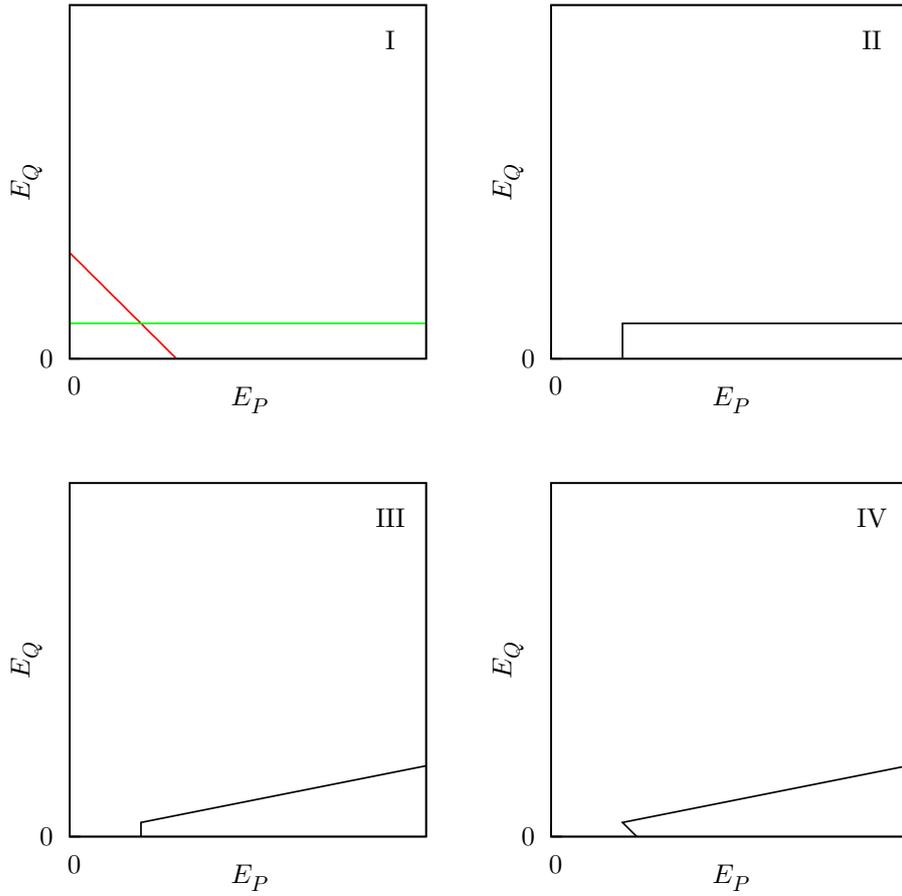


Figure 3.17: I) the PQ plane. The red line roughly bounds the noise region while the green one marks the gw region. II–IV) the gw acceptance regions discussed in the text.

(figure 3.17, box II). However, we expect that such “ideal” region is not convenient in a realistic algorithm. In fact, a practical search necessarily involves a discretization of the sky into a finite set of directions and thus misses the exact gw source direction with high probability. The test direction comes close to the true one, but the error is in practice always finite. On the one hand, this leads to an imperfect synchronization of the time delays; on the other hand, the resulting \mathbf{P} and \mathbf{Q} projectors are not perfectly matched to the gw plane. It follows that there is a slight energy deficiency in E_P and, at the same time, a non-null residual energy ΔE in E_Q . Putting a hard threshold T_2 on E_Q may therefore prevent an efficient detection. Moreover, contrary to intuition, an *intense* gw burst may be easily missed, because— for constant error in the source direction— ΔE is clearly proportional to the total signal energy. The same issue may happen due to calibration errors in the detector transfer functions, because we may expect that they prevent the perfect cancellation of the gw in the Q subspace. The problem was identified in a similar gw discrimination strategy [61] and the proposed solution was to consider the *relative* energy cancellation in the P and Q subspaces, rather than requiring the exact cancellation in E_Q . Indeed, the proportionality between ΔE and the total energy suggests proportionality also between ΔE and E_P . A more robust gw region then resembles a trapezoid, bounded at the top by the line defined as $E_Q/E_P = T_2$:

$$R_2 = \{(E_P, E_Q) \mid E_P > T_1, E_Q/E_P < T_2\} \quad (3.43)$$

(figure 3.17, box III). A slightly different alternative is to use the T_1 threshold on the *total* energy $E_P + E_Q$ and to retain the sloped top edge of R_2 . The resulting region

$$R_3 = \{(E_P, E_Q) \mid E_P + E_Q > T_1, E_Q/E_P < T_2\} \quad (3.44)$$

closely represents a total power detection to distinguish \mathcal{H}_{gw} and \mathcal{H}_s from \mathcal{H}_0 , followed by a discrimination through the statistic E_Q/E_P to distinguish \mathcal{H}_{gw} from \mathcal{H}_s (figure 3.17, box IV).

Once the gw acceptance region R is defined, we can evaluate the alarm probability for a single coefficient of the E_P and E_Q maps by means of the integral

$$p = \int_R P(E_P, E_Q) dE_P dE_Q = \int_R P(E_P)P(E_Q) dE_P dE_Q. \quad (3.45)$$

By using the central χ^2 distributions for $P(E_P)$ and $P(E_Q)$, this gives the false alarm probability, while by using the non-central χ^2 distributions we get the detection probability for gw or spurious signal, depending on the non-centralities.

Thanks to the factorization of $P(E_P, E_Q)$, the case with region R_1 is simply

$$p(T_1, T_2) = \int_{T_1}^{\infty} P(E_P) dE_P \int_0^{T_2} P(E_Q) dE_Q. \quad (3.46)$$

The false alarm probability for Gaussian noise with unit variance can then be obtained by means of the well-known central χ^2 cumulative distribution

functions for 2 and $M - 2$ degrees of freedom [75], i.e.

$$p_{fa}(T_1, T_2) = \left[1 - \frac{\gamma(1, T_1/2)}{\Gamma(1)} \right] \left[\frac{\gamma(M/2 - 1, T_2/2)}{\Gamma(M/2 - 1)} \right] \quad (3.47)$$

where $\gamma(a, b)$ and $\Gamma(a)$ are the lower incomplete and complete Euler gamma functions. In presence of a signal, the non-centralities $s_p = \mathbf{s}^T \mathbf{P} \mathbf{s}$ and $s_q = \mathbf{s}^T \mathbf{Q} \mathbf{s}$ appear: the resulting cumulative distributions for the non-central χ^2 can only be expressed by series expansion. Nevertheless, we can use the second-moment approximation [56] and write the detection probability as

$$p_d(T_1, T_2) = \left[1 - \frac{\gamma\left(\frac{(s_p+2)^2}{4(s_p+1)}, \frac{(s_p+2)T_1}{4(s_p+1)}\right)}{\Gamma\left(\frac{(s_p+2)^2}{4(s_p+1)}\right)} \right] \cdot \left[\frac{\gamma\left(\frac{(s_q+M-2)^2}{2(2s_q+M-2)}, \frac{(s_q+M-2)T_2}{2(2s_q+M-2)}\right)}{\Gamma\left(\frac{(s_q+M-2)^2}{2(2s_q+M-2)}\right)} \right] \quad (3.48)$$

For instance, the probability of detecting a gw burst by testing the correct source direction can be obtained by eq. 3.48 with $s_q = 0$. Note that eq. 3.48 correctly reduces to $p_{fa}(T_1, T_2)$ for $s_p = s_q = 0$.

Using the region R_2 yields instead the alarm probability

$$p(T_1, T_2) = \int_{T_1}^{\infty} P(E_P) \int_0^{T_2 E_P} P(E_Q) dE_P dE_Q \quad (3.49)$$

which can be hardly simplified even for the false alarms produced by Gaussian noise, as the roles of T_1 and T_2 become entangled and we need to integrate a χ^2 cumulative distribution function. Similar complications also hold for region R_3 .

The above calculations account for testing a single coefficient against the gw acceptance region of the PQ plane. Accordingly, they may not match the actual false alarm and detection probabilities for the whole analysis pipeline, which entail searching the whole celestial sphere and time-frequency maps at different levels. The full evaluation of the analysis performances requires complete Monte Carlo simulations, as discussed in the next chapter.

We conclude with the reminder that, as it is well known, current interferometers produce less than ideal data, with noise that may significantly deviate from the Gaussian distribution. Nevertheless, we can construct the virtual channels E_P, E_Q in any case and represent them in the PQ plane. Although it is reasonable to expect the presence of a Gaussian bulk in the noise region, real data will populate this plane differently from what we expect and the optimal gw region could be difficult to define analytically. However, it may be found empirically by constructing the null data sets, e.g. through unphysical time shifts or data surrogation, and by studying their distribution in the PQ plane, in order to fix the acceptable false alarm rate prior to searching for candidate gw bursts. Establishing the effectiveness of the method with realistic noise is another operation that can be performed by means of complete Monte Carlo simulations.

Chapter 4

Implementation and Monte Carlo simulations

Although the statistical behaviour of E_P and E_Q are known analytically for each pixel of the time-frequency plane, it is difficult to accurately predict the combined effects of the time delays $\tau_m(\vartheta, \varphi)$, the wavelet transform and the chosen gw acceptance region in the PQ plane. In particular, we need to assess the detection and discrimination effectiveness of the full method.

A Monte Carlo simulation can be used to fully characterize the overall performances. In fact, we can simulate the arrival of gw bursts with arbitrary waveforms and source direction, evaluate the noisy network response to them and feed such response to a prototype of the analysis algorithm. By carrying out many repeated simulations for each value of the threshold we are then able to estimate both the false alarm probability (by injecting null gw bursts) and the detection efficiency. By combining the two estimates, the ROC curves can thus be evaluated and we can discuss the performances of the method.

Another useful feature of the Monte Carlo method is its ability to use non-Gaussian noise (and possibly even the real detector noise) or to introduce artificial calibration errors. Thus, we are able to test the performance degradation due to such issues.

An obvious drawback of Monte Carlo simulations is that, as many trials must be performed to collect enough statistics, there is need for computing power. This is especially true for network gw data analysis.

4.1 Prototype of the analysis algorithm

We implemented a prototype of the analysis algorithm using the MATLAB environment. Accordingly to what defined in chapter 3, after calculating the M time series $x_m[j]$ representing the detector outputs, the algorithm performs the following operations.

1. A test direction is picked from a suitable mesh on the celestial sphere.
2. P and Q are evaluated for the chosen direction.

3. The series $x_m[j]$ are correctly time-shifted to synchronize the arrival times of a possible gw signal.
4. The M WP decompositions are performed up to a maximum level K . We use our own optimized implementation of the WP decomposition in order to save computing time.
5. The resulting $3M$ -dimensional array of WP coefficients is projected with \mathbf{P} and \mathbf{Q} .
6. The statistics E_P and E_Q are formed.
7. The resulting pair of 3-dimensional arrays is scanned, looking for items falling within the gw acceptance region of the PQ plane. Items falling within the region trigger a gw detection.
8. Operations 1-7 are repeated for another point of the sky, until the whole mesh has been searched or a detection has been recorded.

In order to produce accurate estimates of the alarm rates, the operations are repeatedly ran over many realizations of the network response.

Some of the operations, e.g. step 2, are actually pre-calculated to save computing time.

4.2 Simulation parameters

4.2.1 Detector networks and output signals

In order to verify the effect of the network figures of merit $\|\mathbf{F}\|$ and $\text{cond}(\mathbf{F})$, we compare the effectiveness of the analysis when applied to two very different networks. In particular, we expect that the main parameter controlling the detection efficiency is the “total observed signal strength”

$$h_{\text{obs}} = \left[\int_{-\infty}^{+\infty} \|\mathbf{F} \mathbf{h}(t)\|^2 dt \right]^{\frac{1}{2}} \quad (4.1)$$

which clearly depends on the network figures of merit and on the intrinsic gw signal strength h_{rss} . We thus use the realistic 3-detector LIGO-VIRGO network and the ideal dodecahedral network.

The networks are constructed by using the parameters discussed in chapter 3. The sensitivity curves are assumed identical for every detector, infinite outside the 70–1000 Hz band and constant inside the same band, which is a reasonable approximation to realistic interferometric sensitivity curves of figure 2.1.

The simulated network output signals are time series lasting ~ 0.5 s, oversampled to 10 kHz, populated with white Gaussian noise and band-passed within the 70–1000 Hz band. The total noise variance after band-passing is normalized to 1. The actually analyzed window of samples is slightly reduced with respect to the original ~ 0.5 s due to the time delays $\tau_m(\vartheta, \varphi)$.

The maximum WP decomposition level is set to $K = 6$.

4.2.2 Discretization of the sky

As the distribution of gw sources is not known, in general there is no physical motivation to prefer one sampling of the celestial sphere to another. A reasonable choice is a uniform sampling, e.g. one may choose a uniform grid on ϑ and $\varphi \sin(\vartheta)$ [61].

A second parameter is the number of points, which should be tuned accurately because it determines the computational needs and it clearly impacts the false alarms and detection efficiency. To fix it, we should estimate the characteristic angular scale of the E_P and E_Q statistics in presence of gw bursts. Following eq. 3.10, the angular scale is determined both by the angular factor (i.e. the antenna patterns) and by the geographical factor (i.e. the time delays) of the network response matrix. The former bears a constant, coarse angular scale characteristic of the quadrupolar antenna pattern, resembling the shape of $\cos^2(\vartheta)$. Instead, the latter introduces an angular scale determined both by the detector relative distances (baselines) and by the time scale of the gw signals. In principle, the two scales compete in a complicated way to determine the actual one. In fact, for co-located detectors or extremely slow signals the relative delays in the arrival times are negligible and the angular scale reduces to the antenna pattern one. Instead, for distant detectors or high-frequency signals, the interference fringes caused by the time delays are much finer than the antenna patterns. Nevertheless, the time scale of gw bursts is constrained by the finite bandwidth of the detectors and it can be shown that, for realistic bandwidths and actual detector baselines, the resulting angular scale is much finer than the antenna pattern one (as an example, look at the rather fine angular scale of the interference fringes in figure 3.11).

The detector bandwidth is connected to the frequency at which the data time series are sampled. Then, we intuitively expect that a connection should exist as well between how we choose the sky discretization and the sampling rate of the detector outputs.

Following these considerations, we choose an “automatic” discretization by finding the directions for which the time delays $\tau_m(\vartheta, \varphi)$ are multiple of the sampling time t_s . In other words, we solve the system

$$-\frac{\mathbf{r}_m \cdot \mathbf{k}(\vartheta, \varphi)}{c t_s} = j_m, \quad j_m \in \mathbb{Z} \quad (4.2)$$

where ϑ, φ are the directions to determine. Clearly, larger detector bandwidths imply lower sampling periods which leads to more solutions for system 4.2 and thus to a finer discretization. Moreover, in the case of asymmetric networks like LIGO-VIRGO, the sky is sampled non-uniformly to reflect the actual anisotropic angular resolution of the network.

A minor, computational advantage of this choice is also the fact that the time-shifts performed by the analysis algorithm (step 3) can be implemented as simple shifts in the items of the arrays $x_m[j]$, thus avoiding expensive interpolations.

We solve system 4.2 for each network by means of a preliminary dedicated Monte Carlo algorithm. Random directions are uniformly extracted on the

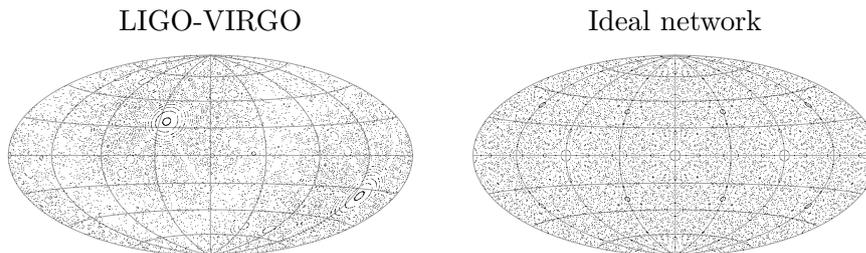


Figure 4.1: discretization of the celestial sphere obtained by numerical solution of system 4.2.

sphere and the associated time delays $\tau_m(\vartheta, \varphi)$ are evaluated, in units of samples of the time series. The algorithm retains only those directions that produce delays with fractional part smaller than a certain maximum error. The sets of the resulting sky samples are sketched in figure 4.1. Interestingly, the LIGO-VIRGO network exhibits an undersampled circular band, which matches the great circle corresponding to the baseline between the two LIGO instruments. The lower sampling can be understood by noting that this baseline is shorter than the others. The consequent lower angular resolution in such regions is automatically accounted for by the lower sampling.

It is worth noticing that the detector outputs are actually oversampled with respect to the ~ 1 kHz natural bandwidth assumed for the detectors, which implies that the sky is correctly oversampled as well.

We gathered $\sim 10^4$ solutions of system 4.2 for both networks, but we are only using about $2 \cdot 10^3$ of them to relax the computational requirements.

4.2.3 Burst waveforms

Recalling the discussion of gw bursts in chapter 1, a problematic choice of any Monte Carlo simulation involving gw bursts is our inadequate knowledge of the actual manifold of gw waveforms. In fact, the resulting estimated performances of the simulated detection algorithms have a meaning only with respect to the tested signals.

Usually, gw detection algorithms are tested with simulated waveforms from numerical relativity, as in [61], or with generic oscillating waveforms like sine-Gaussians and transients with central frequency sweeping in time (*chirplets*) as in [54].

For simplicity we adopt sine-Gaussians, i.e. the injected signals have the form

$$x(t) = A \exp \left[- \left(\frac{t - t_0}{\Delta t} \right)^2 \right] \cos(2\pi f t + \phi) \quad (4.3)$$

where A is the amplitude, t_0 is the arrival time, Δt is the duration, f is the central frequency and ϕ is the initial phase. t_0 is placed at the center of each simulated data set and a uniformly random jitter (within a few samples) is added to it, in order to simulate the unknown arrival time and avoid possible issues due to the synchronization between the signal and the sampling of the

data set. ϕ is simply extracted from a uniform distribution in the range $0-2\pi$. Δt and f are the most important parameters because they control how much of the signal energy—in the sense of h_{rss} —overlaps with the limited detector bandwidth. Moreover, they control the time-frequency extension and location and thus determine how the signal behaves with respect to the WP decomposition. Δt is chosen in the range 10^{-3} – 10^{-2} s in order to keep the signal magnitude negligible at the extremes of each data set and avoid border issues due to truncation. Similarly, f is chosen to avoid crossing the 70–1000 Hz bandwidth boundary.

The two waveforms associated with the two polarization states share the same parameter values, except for the initial phase ϕ which is extracted separately for each polarization.

4.3 Results of the Monte Carlo simulations

The simulations are currently being performed thanks to the computing resources provided by the AURIGA group. The analysis of a single data set with the presented values of the parameters takes 100–200 s on single-core machines with ~ 2 GHz clock speeds.

The preliminary results involve the estimation of the false alarm probability p_{fa} for different extensions of the R_3 acceptance region defined in chapter 4 and the estimation of the detection efficiency p_d for intense gw bursts (i.e. $h_{\text{obs}} \simeq 10^2$). The observed dependency of p_{fa} on the threshold values bears the characteristic exponential decay we expect from χ^2 distributions, particularly with T_1 which thresholds the total energy $E_P + E_Q$. The preliminary estimate of p_d for intense gw bursts is indistinguishable from 1 for all the tested threshold values, which makes us confident about the Monte Carlo implementation.

The next step to be performed is a more accurate estimation of p_{fa} and the estimation of p_d for lower-energy gw bursts, which completes the estimation of the ROC curve around realistic and interesting values of p_{fa} , e.g. 10^{-4} .

Chapter 5

Conclusions and future perspectives

This thesis deals with the definition of a procedure for detecting gw transients out of the detector noise and for discriminating them from instrumental, non-gravitational transients, which represent a major challenge in current gw searches. The main innovations introduced with this work are the exploitation of the gw physical signatures directly in the detection process, rather than using them as post-detection vetoes. In addition, detection is accomplished by a suitable acceptance region of the virtual PQ plane.

First we defined the generic scheme of an analysis pipeline for searching gws bursts with a network of detectors. We suggested the synthesis of “virtual channels” to recognize the physical properties of gws. In order to tackle the detection and discrimination problem, the network response to gws and the gw inverse problem were then reviewed. This led us to discuss a number of network figures of merit, in particular $\|\mathbf{F}\|$ and $\text{cond}(\mathbf{F})$, that determine the effectiveness of a network for performing gw searches. The performances of actual networks were studied under these figures of merit and we pointed out that optimal or very good networks are in principle feasible. We then evidenced that a simple geometrical interpretation can be given to the network response, we recognized the physical signatures of gws within such interpretation and we proposed to exploit such property for the detection and discrimination of genuine gw bursts. A convenient way is through geometric projections of the measured data. We showed how to perform such projections and arrived at the definition of the synthetic channels E_P and E_Q . We proved, both analytically and with simulations, that under the assumption of Gaussian noise their statistics are simple and well-behaved, even for non-ideal and ill-conditioned networks. The E_P , E_Q channels were used to define the virtual PQ plane, which is populated differently depending on the kind of signal contained in the measured data, thanks to the statistics of E_P and E_Q . Accordingly, we moved the detection and discrimination problem to the PQ plane by tracing a gw acceptance region R in the plane and checking whether the network data enters R for some directions in the sky. We accounted for the unknown time-frequency structure of gw bursts by exploiting the multiresolution analysis provided by

an orthogonal wavelet packet decomposition, performed at multiple levels.

Finally, we implemented a prototype of the detection/discrimination algorithm and we ran complete Monte Carlo simulations to characterize its performances.

Future work entails verifying the expected dependence of the detection efficiency on the network figures of merit and testing the algorithm with a wider class of transient signals—e.g. chirplets—to study the variation in detection efficiency over the signal manifold accessible to realistic gw detectors. Clearly, the final goal is the validation of the method with respect to real data from current detectors, e.g. LIGO and VIRGO. In particular, the study of how the real data populates the PQ plane is mandatory.

Furthermore, we believe that the geometrical interpretation of the network response, and in particular the planar structure of gws, can be used for the definition of other detection statistics, and we are investigating such research lines. In principle, the method is applicable not only to impulsive gw searches but actually to any gw signal. However, different detection statistics could be more suited. Another interesting extension should be the identification of more sources in the sky that contribute to the same data set.

Bibliography

- [1] S. Weinberg, *Gravitation and cosmology*, John Wiley & Sons
- [2] C. W. Misner, K. S. Thorne, J. A. Wheeler, *Gravitation*, W. H. Freeman and Company (1973)
- [3] Taylor et al., *Measurements of general relativistic effects in the binary pulsar PSR1913+16*, J.M. 1979, Nature 277, 437
- [4] M. Kramer et al., *Tests of general relativity from timing the double pulsar*, Science, Volume 314, Issue 5796, pp. 97-102 (2006)
- [5] C. Cutler, K. S. Thorne, *An overview of gravitational-wave sources*, General Relativity and Gravitation, Proceedings of the 16th International Conference, ed. N. Bishop and S. D. Maharaj (World Scientific, 2002)
- [6] C. Van Den Broeck, *The gravitational wave spectrum of non-axisymmetric, freely precessing neutron stars*, Class. Quant. Grav. 22 1825-1840 (2005)
- [7] H. Dimmelmeier et al., *Generic gravitational wave signals from the collapse of rotating stellar cores*, Phys. Rev. Lett. 98, 251101 (2007)
- [8] M. Kawamura, K. Oohara, T. Nakamura, *General relativistic numerical simulation on coalescing binary neutron stars and gauge-invariant gravitational wave extraction*, astro-ph/0306481 (2003)
- [9] M. Shibata, K. Taniguchi, *Merger of black hole and neutron star in general relativity: tidal disruption, torus mass, and gravitational waves*, Phys. Rev. D 77, 084015 (2008)
- [10] M. Campanelli, C. O. Lousto, Y. Zlochower, *Last orbit of binary black holes*, Phys. Rev. D 73, 061501(R) (2006)
- [11] M. Maggiore, *Gravitational waves and fundamental physics*, gr-qc/0602057v1
- [12] B. Allen, *The stochastic gravity-wave background: sources and detection*, gr-qc/9604033v3 (1996)
- [13] X. Siemens, V. Mandic, J. Creighton, *Gravitational wave stochastic background from cosmic strings*, Phys. Rev. Lett. 98, 111101 (2007)

- [14] P. R. Saulson, *Physics of gravitational wave detection: resonant and interferometric detectors* (1998)
- [15] LIGO, <http://www.ligo.caltech.edu>
- [16] VIRGO, <http://www.virgo.infn.it>
- [17] GEO600, <http://geo600.aei.mpg.de>
- [18] TAMA300, <http://tamago.mtk.nao.ac.jp>
- [19] AURIGA, <http://www.auriga.lnl.infn.it>
- [20] EXPLORER, http://www.lnf.infn.it/esperimenti/rog/frame_explorer.htm
- [21] NAUTILUS, http://www.lnf.infn.it/esperimenti/rog/frame_nautilus.htm
- [22] International Gravitational Event Collaboration, <http://igec.lnl.infn.it>
- [23] IGEC-2 Collaboration, *Results of the IGEC-2 search for gravitational wave bursts during 2005*, Phys. Rev. D 76, 102001 (2007)
- [24] LIGO Scientific Collaboration, *Advanced LIGO reference design*, LIGO-060056-08-M
- [25] A. Spallicci et al. *Advanced VIRGO: detector optimization for gravitational waves by inspiralling binaries*, Class. Quant. Grav. 22 S461-S469 (2005)
- [26] L. Gottardi, *Complete model of a spherical gravitational wave detector with capacitive transducers. Calibration and sensitivity optimization*, Phys. Rev. D 75, 022002 (2007)
- [27] DUAL, <http://dual.lnl.infn.it>
- [28] M. Cerdonio, *Acoustic GW detectors in the 2010 timeframe*, Class. Quant. Grav. 20 1-7 (2003)
- [29] Large Interferometer Space Antenna, <http://lisa.jpl.nasa.gov>
- [30] Ray d'Inverno, *Introducing Einstein's relativity*, Oxford University Press
- [31] H. Bondi, *Plane gravitational waves in general relativity*, Nature 179 (1957)
- [32] S. W. Hawking, W. Israel, *Three hundred years of gravitation*, Cambridge University Press (1987)
- [33] A. Einstein, *The field equations of gravitation*, Preussische Akademie der Wissenschaften, Sitzungsberichte (1915) (part 2)
- [34] C. M. Will, *The confrontation between general relativity and experiment in Living Reviews in relativity* (lrr-2006-3) (2006)

- [35] N. Ashby, *Relativity in the Global Positioning System* in *Living Reviews in relativity* (lrr-2003-1) (2003)
- [36] H. Lichtenegger, B. Mashhoon, *Mach's principle*, physics/0407078v2
- [37] J. C. Baez, E. F. Bunn, *The meaning of Einstein's equation*, Amer. Jour. Phys. 73, 644-652 (2005)
- [38] Gel'fand, Minlos, Shapiro, *Representations of the rotation and Lorentz groups*, Pergamon Press, London (1963)
- [39] N. Arnaud, M. Varvella, *Angular patterns of interferometric detectors and resonant bars for tensorial and scalar gravitational waves*, LAL 04-115
- [40] J. Weber, *Gravitational-wave-detector events*, Phys. Rev. Lett. 20, 1307 - 1308 (1968)
- [41] R. L. Forward, *Wideband laser-interferometer gravitational-radiation experiment*, Phys. Rev. D 17, 379 - 390 (1978)
- [42] B. Abbott et al. (LSC), *Search for gravitational wave bursts in LIGO's third science run*, Class. Quant. Grav. 23 S29-S39 (2006)
- [43] B. Abbott et al. (LSC), *Upper limits on a stochastic background of gravitational waves*, Phys. Rev. Lett. 95, 221101 (2005)
- [44] LIGO Scientific Collaboration, *The Einstein@Home search for periodic gravitational waves in LIGO S4 data*, LIGO-P080021-01-Z
- [45] A. M. Sintes, *Report on an all-sky LIGO search for periodic gravitational waves in the S4 data*, LIGO-P070121
- [46] *Gravitational waveform catalog*, http://www.mpa-garching.mpg.de/rel_hydro/wave_catalogue.shtml
- [47] P. Leaci et al., *Antenna pattern of DUAL detectors of gravitational waves and its exploitation in a network of advanced interferometers*, submitted to Class. Quant. Grav. (2008)
- [48] A. Dietz et al., *Calibration of the LIGO detectors for S4*, LIGO-T050262-01-D (2006)
- [49] E. Hirose, X. Siemens, *Validation of time-domain calibration*, LIGO-T070275-00-Z (2007)
- [50] L. D. Landau, E. M. Lifshitz, *Statistical physics*, 3rd edition, part 1
- [51] P. Astone et al., *Cosmic rays observed by the resonant gravitational wave detector NAUTILUS*, Phys. Rev. Lett. 84, 14 (2000)
- [52] M.-A. Bizouard et al., *Data quality and veto studies for the GW burst search in Virgo C7 Run data*, VIR-013D-07 (2007)

- [53] LSC Glitch Group, *The LSC glitch group: monitoring noise transients during the fifth LIGO science run*, LIGO-P080016-00-Z (2008)
- [54] M. Camarda, A. Ortolan, *Search for gravitational wave bursts by wavelet packet decomposition: the detection algorithm*, Phys. Rev. D 74, 062001 (2006)
- [55] C. W. Helstrom, *Statistical theory of signal detection*, Pergamon Press, Oxford (1968)
- [56] W. G. Anderson et al., *Excess power statistic for detection of burst sources of gravitational radiation*, Phys. Rev. D 63, 042003 (2001)
- [57] W. G. Anderson, R. Balasubramanian, *Time-frequency detection of gravitational waves*, Phys. Rev. D 60, 102001 (1999)
- [58] S. D. Mohanty, *Robust test for detecting non-stationarity in data from gravitational wave detectors*, Phys. Rev. D 61, 122002 (2000)
- [59] S. Klimenko et al., *Performance of the WaveBurst algorithm on LIGO data*, Class. Quant. Grav. 21 S1685-S1694 (2004)
- [60] Y. Gürsel, M. Tinto, *Near optimal solution to the inverse problem for gravitational-wave bursts*, Phys. Rev. D 40, 3884 (1989)
- [61] S. Chatterji et al., *Coherent network analysis technique for discriminating gravitational-wave bursts from instrumental noise*, LIGO-P060009-01-E (2006)
- [62] L. Cadonati, *Coherent waveform consistency test for LIGO burst candidates*, Class. Quant. Grav. 21, S1695-S1704 (2004)
- [63] M. Rakhmanov, S. Klimenko, *A cross-correlation method for burst searches with networks of misaligned gravitational-wave detectors*, Class. Quant. Grav. 22 S1311-S1320 (2005)
- [64] S. Klimenko et al., *Constraint likelihood analysis for a network of gravitational wave detectors*, Phys. Rev. D 72, 122002 (2005)
- [65] S. Klimenko et al., *A coherent method for detection of gravitational wave bursts*, Class. Quant. Grav. 25, 114029 (2008)
- [66] A. Pai, S. Dhurandhar, S. Bose, *Data-analysis strategy for detecting gravitational-wave signals from inspiraling compact binaries with a network of laser-interferometric detectors*, Phys. Rev. D 64, 042004 (2001)
- [67] See for example the review *Inverse problems in science and engineering*, published by Taylor & Francis
- [68] M. Rakhmanov, *Rank deficiency and Tikhonov regularization in the inverse problem for gravitational-wave bursts*, LIGO-P060005 (2006)
- [69] A. Ben-Israel, T. Greville, *Generalized inverses: theory and applications*

- [70] P. Hansen, *Rank-deficient and discrete ill-posed problems*, Philadelphia: SIAM (1998)
- [71] A. Tikhonov, V. Arsenin, *Solutions of ill-posed problems*, Washington, D.C.: V.H. Winston & Sons (1977)
- [72] P. R. Saulson, *Fundamentals of interferometric gravitational wave detectors*, World Scientific Publishing (1994)
- [73] M. Cerdonio et al., *Concept of a resonant antennæ observatory for gravitational wave bursts*, Phys. Rev. Lett. 71, 4107-4110 (1993)
- [74] S. Mallat, *A wavelet tour of signal processing (second edition)*, San Diego Academic Press (1998)
- [75] A. Papoulis *Probability, random variables, and stochastic processes*, McGraw-Hill, Singapore (1984)

Acknowledgments

First of all, I thank my family and particularly my parents Silvia and Varo. My university experience and my current interest for physics would have been much more difficult to achieve without their support and encouragement. It is a pleasure to devote this work to them, as a little gift for their efforts.

I thank the AURIGA team for their kind hospitality and for providing their computing resources. I am particularly grateful to Antonello Ortolan for the exciting things we have been working on in the last year, for continuously keeping my brain active with new questions and problems and for always respecting my work, let alone for introducing me to the world of the experimental physics and for allowing me to visit several labs across Italy.

My life would be very different without the MeTA~LAbS people. The moments we passed together are among the best ones in my life. After all the adventures we experienced together, our roads are now apparently diverging, but I hope they will join again in the future, in the name of the foobar. I think our deep objectives are the same after all.

I want to thank all the people who lived with me in Padova, particularly who first convinced me to move to that beautiful city. We were extremely different, but we shared many emotional moments. Everyone of them has been the subject of many thoughts of mine and now forms a bit of my personality.

Of course, Silvia & Giorgio deserve a special mention for the invaluable service they have been unconsciously providing. Thank you!