UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica “*Galileo Galilei*”

DOTTORATO DI RICERCA IN FISICA
CICLO XIII

DATA ANALYSIS FOR GRAVITATIONAL WAVE SEARCH WITH RESONANT DETECTORS

Coordinatore: Ch.mo Prof. Attilio Stella
Supervisore: Ch.mo Prof. Massimo Cerdonio

Dottorando: Lucio Baggio

Preface

It was a privilege for me to work in the last four years with the AURIGA group, a team of brilliant, passionate and crazy scientists, who are devoting themselves to a research that someone with common sense can well describe desperate. But how would any pioneer be described, anyway? Day by day, the intense discussions, critics and cheers among us helped to motivate our work, and the willing for an endless improving.

Now I am at a turn, the work is not finished –endless, as I said– but it’s time to have a look backward to what has been done up to now. My original intention, as I can clearly see now, was probably to write a book, so many and so different topics I touched. But time and human resources are limited, so the output is perhaps just a humble compendium of what is the usual subject of daily work in the field of gravitational wave experimental data analysis.

Even if there are probably better lectures on the field, I tried to target this paper to as many readers as possible, first of all by choosing to write it in English (well, at least I tried!). After a brief introduction, I enter the main thread of signal and noise separation, which develops in the first two chapters, with a little tail about multidetector data analysis, because we are not alone in the Earth, are we?. However many interesting specialist (or just too miscellaneus) topics were moved to Chapter 4, which you are not expected to read as the last one, neither progressively.

As to the style of writing, you will see that I often oscillate between the third person (which undoubtedly sounds “objective” and “firm”) and the first plural (“we”), rather than using the singular (“I”), because ultimately I feel in debt for every result –directly or indirectly– with the rest of the team.

Legnaro (PD, Italy), 1 Jan 2001

Lucio Baggio

CONTENTS

Introduction

One of the most important conceptual revolutions in this century is without doubt the idea of a universe with locally curved geometry, which influences and is influenced by mass density distribution. According to Einstein's theory of General Relativity (GR), the Earth doesn't turn around the Sun because of a direct interaction, but only because it feels the curvature in the bulk of space-time induced by the Sun itself. As the curvature field interacts with matter only locally, we need a way for this field to propagate at large distance from the source. This can be done only at speed equal or less than light, which means that this moving field –i.e. gravitational waves (gw)– must survive and be self-sustained after it detaches from the source.

Up to now GR had many important experimental consequences verified to a high degree of accuracy, last but not least the production of gravitational waves by observing angular momentum loss in binary pulsar system. The next challenge in fundamental physics is the direct detection of gravitational waves by an Earth or space based device. First of all, to put another brick on the fabric of experimental validation of GR. Second, to use gravitational waves as a new probe of cosmic sources.

Building a gravitational wave detector is one of the frontiers of experimental physics. Cryogenic resonant gw detectors are the first and, up to now, only devices sensitive to at least the most violent gw emission occurring in the galaxy. Although such events are predicted to be extremely rare, still resonant detectors will continue to give at least relevant upper limits, if not detection, as they operate now and, after upgrades, in the future. Resonant detectors are basically very sensitive oscillators, capable at their "quantum" limit of detecting changes in their vibrational amplitude of the order of less than a millionth of the diameter of an atom nucleus, or 10^{-21}m, which translates (see below) in gw amplitudes of some $h\sim 3\ 10^{-22}$. They are designed to detect gravitational waves through the excitation of the quadrupole resonant modes of massive cylinders or spheres of high mechanical quality factor, cooled at liquid helium temperatures and below to reduce spontaneous noise.

There are now five bar detectors in long term operation, ALLEGRO [5] at Baton Rouge in US, AURIGA[35] at Legnaro in Italy, EXPLORER [3] at CERN, NAUTILUS [4] at Frascati in Italy and NIOBE [1] at Perth in Australia, and since 1997 they have joint in a single international data exchange and coincidence search community, known as IGEC (International Gravitational Event Collaboration). [29][24]

The five detectors in IGEC have rather similar schemes. The sensitive part of these devices is a cylindrical shaped mass of Aluminum or Niobium, few meters long and few tons in weight, with high mechanical ('Q') factor, of the order of ten million. It is mechanically decoupled from the environment noise (seismic, acoustical, and so on), while intrinsic thermal vibration of the crystal lattice is kept as low as possible by reaching temperatures below liquid Helium, in the range 2÷0.1K.

A gw impinging transversally to the bar axis changes the amplitude of the first longitudinal compression mode of oscillation. To detect this signal, an auxiliary oscillator with a mass in the kg range is attached to one of the bar faces, and its resonance frequency tuned to that of the sensitive mode of the bar in order to have a strong privileged coupling. This transducer is in turn electrically coupled (with a variety of solutions: capacitive, inductive, microwave, optical) to the external readout.

Even if the experimental setup are not identical, the sensitivity is similar, within a factor of four in the energy of a detectable ms gw burst, and corresponds to that delivered in a violent emission of >0.01 M-sun in the Galaxy as in the coalescence of neutron star or black hole binaries. In amplitude of the metric perturbation at the detector, the present burst sensitivity is about $h = 5\ 10^{-19}$; the useful bandwidths are of few Hz around the kHz resonant frequencies. A factor of at least 10 in sensitivity is expected, after the upgrades under way.

The five detectors are oriented with their axes parallel one to each other and each locally orthogonal to a Earth great circle, which luckily happens to pass very close to each location, in order to have all their antenna patterns oriented coherently and thus maximize the coincidence probability.

Even when the sensitivity of detectors enters the range of loudest astrophysical sources, the amplitude of detected signals would not be much high. Therefore, the very first important issue with data analysis in gravitational wave detectors is that they are *dominated by noise*. This is a simple but important concept to keep in mind while going through the reading.

A background of sharp and clearly time-limited signals characterizes most of other experiments, especially in particle physics, with signal amplitude vastly above the level of electronic noise. Part of these events originates within the detector itself; part is due to the local environment. The latter ones can be somewhat diminished by appropriate shielding of the detector, and both can be cured in part by anticoincidence or trigger methods, and where possible by other selections based on physics properties of the signal itself (time of flight, electrical charge, mass, spin and so on). Similar techniques can be applied to the gw detection when performing a multi detector analysis. Time of flight consistency for the time of arrivals and tracelessness of the metric tensor are two physical requirements that can be used to gain specific gw signal signatures against the background. [6]

But a signal-to-noise ratio of maybe ten or more is requested, while data exchanged by detectors are composed mainly by candidate gw events with SNR equal or less than five. Where are they coming from? They are random fluctuations of the background Gaussian noise (thermal noise of the bar and of the electromechanical transducer plus white noise of the amplifier), which are not negligible with such a low threshold. As such, they are very fundamental, and there is nothing that can be done to avoid them (apart from having a detector with better sensitivity, of course). In this low SNR region a coincidence analysis suffers, because lowering the threshold pays a little gain in sensitivity with a powerful increase of the false alarm rate. On the other hand, environmental and internal disturbances fill up the higher SNR region, where the present coincidence analysis can do a good cleaning job.

The bulk of the data acquisition and data analysis –from the raw data to selected event candidates– was historically developed independently by the different groups that operate the detectors.

This work will describe the data analysis system for the production of candidate events for gw bursts, focusing on a particular experiment, the AURIGA detector operating at INFN National Laboratories of Legnaro, Italy. This detector joined the community of gw detectors with the final (ultracryogenic) experimental setup on June 1997 [35].

The AURIGA detector consists of a 2.3tons-3 meters long Al5056 bar hanging on a multiple stage pendulum attenuation system (-240db overall at 1kHz). The latter is made by three cylindrical heavy copper thermal shields suspended one inside another by Ti cables, the inner one supporting the bar by a U-shaped copper cable. The bar and the shields are embraced by a large thoroidal liquid He container, and thermally linked to a ^{3}He-^{4}He-dilution refrigerator that traverse the shields and keeps the bar at 0.2K. The chief features of the outer part of the cryostat are Al-Mylar foils thermal insulations and a high vacuum steel container. Lead-rubber piles at room temperature complete the attenuation system.

A resonant "mushroom" capacitive transducer [8] (0.4kg, 3nF) is coupled to the first longitudinal mode of the bar (920 Hz) and charged with a ~10^6V/m electric field. The electrical transduction chain is made by a superconductive transformer and a commercial dc-SQUID preamplifier equipped with feedback circuit to linearize its output. An high linearity (23-bit at about 4.9 kHz) ADC acquires the signal channel after the amplifier. The time tag and synchronization with UTC of the acquired data is achieved in hardware well within 1μs by dedicated interrupts and triggers between the ADC and a GPS clock with a stabilized local oscillator. The full raw data are then fully archived and analyzed on-line. [36]

Beside the actual implementations, the task to accomplish from this point forward is to enhance and detect a rare gw burst that hopefully could be revealed with an amplitude signal to noise ratio (SNR) as small as 3 (i.e. 10dB). The online analysis system was developed with a stationary and Gaussian noise in mind, so its core is a standard non-adaptive Wiener-Kolmogorov (WK) linear filter, matched to a δ-like (i.e. of millisecond duration) gw burst impinging on the bar. This scheme was slightly modified to track slow drifts in the parameters needed to define the WK filter.

The WK filter is implemented in AURIGA data analysis as a parametric ARMA model, therefore, granted that the data are enough clean to be fitted by this simple model, its 9 fundamental parameters can be computed in a deterministic way by the same number of integrated quantities derived from the power spectral density (PSD) estimates. The WK filter uses each hour a new set of parameters; therefore slow drifts (like the change in the mode frequency due to discharge in the capacitive transducer) are correctly handled. The mode frequencies are followed with digital lock-ins in the raw data, while the quality factors of the modes are just measured once per acquisition run, as they depend on major setup parameters of the detector. The bandwidth of the WK filter is adaptively corrected by a feedback on the residual ‘color’ around the modes in the whitened data PSD.

While the theory behind optimum filtering is well understood, validation of results in presence of a non-stationary system is not as straightforward. A subjective decision has to be taken to state that our knowledge of the behavior of the system is good or poor.

We have chosen as a rule of the thumb that for the system to be considered in a quasi-stationary regime, the typical time of variation should be at least of the order of the thermal/mechanical relaxation time. Every other transient is non-modeled, and are tolerable as long as they do not produce too many false alarms. A big short-lived excitation that enters the system –either gw signal or spurious– spoils the whitened noise PSD estimate, but also the histogram of WK filtered data, where huge tails appear outside the bulk of Gaussian samples. We take care of this by freezing the parameters estimation when such non-Gaussian behavior is detected.

An important issue about this is that entering a long period of non Gaussian noise means “to drive in the fog”, and eventually –if the freezing has lasted a time of the order of the drifting time of the parameters, the information on the noise properties are not considered reliable any more, and the results of the entire period is discarded.

It is useful to remark that this coarse solution is probably not the optimum in terms of duty cycle. Yet it has the merit of being automatic, deterministic, and need only a fast manual check by a supervisor, mainly to take care of ambiguous situations –e.g. a perfectly “clean” Gaussian noise when the SQUID amplifier is locked in a fault position. In particular we tried to avoid any *a posteriori* kind of selection of the data.

As well as noise filtering, event search is model dependent as well. A candidate δ-like event produces a pattern in the WK filter output with a specific mix of an exponential decay, a beat modulation between the two modes and a carrier wave. A high precision max-hold algorithm determines the amplitude of the event, and the decorrelation delay (i.e. the dead time) of the algorithm is at least the time constant of the WK filter, to make sure that the event is not counted twice. Once the event parameters are determined, a goodness-of-the-fit test checks the compliance of the event with the template, and eventually accepts or rejects it.

The detection of an event in a single detector is not sufficient to claim any discovery with high confidence. That's why at least a two-detector coincidence search must be performed (and this also costs in terms of the duty cycle, unless there is a strong effort to coordinate the operative time of the two detectors). One of the main tasks in multi-detector data analysis dealing with event lists is how to correctly compute the false alarm/false dismissal probabilities in order to compute an upper limit for the incoming amplitude and rate of gravitational wave bursts (as it is unlike that any spectacular high signal-to-noise ratio triple coincidence is going to show up with the present configuration of the IGEC observatory).

The main issue within the IGEC is to coordinate the multi-detector data analysis, and with the prospect of being capable in the future of spreading news about highly significant claims of gw detection with a fast alarm to all interested scientists in the worldwide community.

AURIGA joined the community of data taking detectors in June 1997 [35], and one month after I had my final graduating examination. In the successive three years I was involved in the big effort of finishing the complete test of AURIGA data analysis, whose successive advanced versions were month-by-month alternating in the online data analysis. In the mean time the IGEC (signed in the same June 1997) was urging a protocol for data exchange, that started slowly, and became operative only in late 1998, with the first long term data exchange between 5 detectors. As a member of the IGEC data analysis *task force*, I contributed to the first version of the data exchange protocol, and I was in charge of its successive critical revision along with G.A. Prodi, with whom I provided a theoretical framework for the coincidence search data analysis. We eventually proposed a more mature second version of the protocol, which after a long discussion was finally approved a few weeks ago for use in the IGEC future data exchange, which is going to restart soon.

In this work I am going to give a fast survey of all the topics about signal and noise separation, data validation and conditioning, multi-detector data analysis and exchange to which I actively contributed.

Chapter 1 covers the main theoretical thread, the standard signal analysis for Gaussian stochastic noise. When it was possible, a fast derivation of each formula was given, so that the text results more self-contained.

Chapter 2 tries to give a round look to the way this approach does (and sometimes does not) apply to real data of the detector. Chapter 3 covers a few topics about multi-detector analysis, in particular how to deal with the estimate of false alarms and false dismissals.

Chapter 4 covers a number of separate bench tests and simulations to check the various aspects of the methods and describes in detail a few specific implementations in AURIGA. One among the latter is a recent development of AURIGA toolset –an online detection efficiency estimator by Monte Carlo techniques– which can prove useful in single as well in multi detector methods.

The gw strain amplitude is expressed in terms of the metric perturbation h(t) (adimensional units), which represents the relative variation in length of measurements in a curved space-time; for a pair of free falling masses at distance L, the effect of the gw is to change L by ΔL such that h~ΔL/L. The Fourier spectral component of the gw amplitude is $H(\omega)$ (units of Hz^{-1}). For a gw burst we take $h(t) = H_0\delta(t)$. To characterize the noise of the detector, it is usually expressed in terms of the spectral power $S_{hh}(\omega)$ (units of Hz^{-1}) of the equivalent gw amplitude noise h(t), which – through the transfer function of the system – is converted to the output noise power spectrum of the detector (e.g. [34]).

1 Optimal filtering for well behaved systems

1.1 Optimal linear Wiener-Kolmogorov filter

A convenient set of hypothesis to build up a consistent filter theory is that the signal is modified only through linear systems and additive Gaussian noise, the noise itself is stationary and its autocorrelation matrix is known. Under these conditions the Wiener-Kolmogorov (WK) filter is the minimum-variance unbiased linear estimator of the signal amplitude.

Lacking knowledge of the time of arrival, the WK filter is applied continuously to the stream of incoming data, and thus it resembles a noncausal time-invariant linear system. An event search algorithm is then required to locate the signal in the filtered data.

1.1.1 Introduction: signal and noise

The output of any measurement device, or *detector*, is a (possibly digitally recorded) *signal* made by two mixed components, one from the physical phenomenon under study, the other from different sources –even internal to the detector itself. It is customary to refer to the latter as *noise*, and to the first just as *the signal*. The purpose of data analysis is, in short, to separate the signal from the noise, keeping the first and throwing the second. In fact, this very complicated task can be afforded only after some simplifying assumptions –or a model for the detector properties– have been chosen.

In this chapter, we shall restrict ourselves to the following very powerful hypothesis:

A. The detector is a *linear stationary system*;
B. The noise is a *stationary stochastic process with zero-mean Gaussian*[1] *statistics* that linearly superimpose to the signal.

Let us make these statements more precise by introducing some notation and definitions. Unless stated otherwise, a *linear system* denoted by G is described by a *transfer function* $\mathrm{G}(\omega)$, or equivalently by its *impulse response* $\mathrm{g}(t)$. The notation $G \circ \mathrm{f}(t)$ represents the function $\mathrm{f}(t)$ composed with the linear system G:

$$G \circ \mathrm{f}(t) \equiv \int_{-\infty}^{+\infty} \mathrm{g}(\tau)\,\mathrm{f}(t-\tau)\,dt \quad \Leftrightarrow \quad G \circ \tilde{\mathrm{f}}(\omega) \equiv \mathrm{G}(\omega)\,\tilde{\mathrm{f}}(\omega) \tag{1.1}$$

where the Fourier transform (FT) is hereafter defined as

$$\tilde{\mathrm{f}}(\omega) \equiv \int_{-\infty}^{+\infty} \mathrm{f}(t) e^{i\omega t} dt \quad \Leftrightarrow \quad \mathrm{f}(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \tilde{\mathrm{f}}(\omega) e^{-i\omega t} d\omega \tag{1.2}$$

A *random variable* (RV) $\mathbf{x}$ takes its values in the set of all possible outcomes of a measurement, and it is described by its *(first-order) distribution* $F_x(x) = P\{\mathbf{x} \le x\}$ (i.e. the probability of the event $\{\mathbf{x} \le x\}$) or equivalently by its derivative, the density function $f_{\mathbf{x}}(x)$.

A Gaussian statistics is characterized by

$$f_{\mathbf{x}}(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{(x-\bar{x})^2}{2\sigma_x^2}} \tag{1.3}$$

---

[1] It is the kind of statistic one expects if as the final result of a very large number of independent contributions of the same order of magnitude (*central limit theorem*, e.g. [11]).

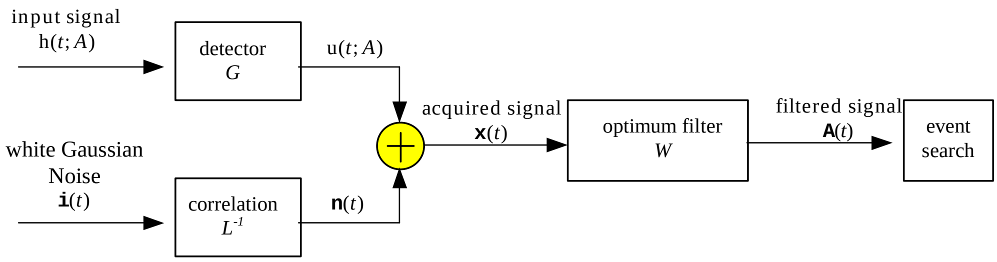


Figure 1 – Schematic of the detector from a point of view of the data analysis system. The boxes represent linear systems, the "+" means linear superposition. The digital filter *W* is the last linear step in the analysis, and it is followed by a non-linear threshold-crossing search algorithm (see 1.1.3).

where $\bar{\mathbf{x}} \equiv \langle \mathbf{x} \rangle \equiv E\{\mathbf{x}\} \equiv \int_{-\infty}^{+\infty} x f_{\mathbf{x}}(x)dx$ is the *mean* of the RV x, and $\sigma_{\mathbf{x}}^2 \equiv Var(\mathbf{x}) = \left\langle (\mathbf{x}-\bar{\mathbf{x}})^2 \right\rangle$ is its *variance*.

A *stochastic process* is a set $\mathbf{x}(t)$ of RV's tagged by the time variable t. The correlation properties of $\mathbf{x}(t)$ at different times are described either by the *autocovariance* $R_{\mathbf{nn}}(\tau)$ or by the corresponding *spectral density* $S_{\mathbf{nn}}(\omega)$:

$$R_{\mathbf{nn}}(\tau) \equiv \left\langle (\mathbf{n}(t)-\bar{\mathbf{n}})(\mathbf{n}(t+\tau)-\bar{\mathbf{n}}) \right\rangle \tag{1.4}$$

$$S_{\mathbf{nn}}(\omega) \equiv \int_{-\infty}^{+\infty} R_{\mathbf{nn}}(\tau) e^{i\omega\tau} d\tau \tag{1.5}$$

where we are assuming that the process $\mathbf{n}(t)$ is *stationary* (hence $R_{\mathbf{nn}}(\tau)$ does not depend on absolute time coordinate).

It can be easily derived from (1.5) that the spectral density of the stochastic process $\mathbf{n}$ after the linear system T is

$$S_{T\circ\mathbf{n}T\circ\mathbf{n}}(\omega) = |\mathrm{T}(\omega)|^2 S_{\mathbf{nn}}(\omega) \tag{1.6}$$

We can restate our working hypothesis. Let $\mathrm{f}(t)$ be the input signal, T the linear system modeling the detector response to an excitation, $\mathbf{n}(t)$ the noise ($\langle \mathbf{n}(t) \rangle = 0$). An interesting consequence of linearity is that the net effect of the different noise sources acting at different ports of the detector is exactly equivalent to having a noise signal added immediately before the digital processing box (see Figure 1).

Therefore, the acquired signal $x(t)$ can be written as

$$\mathbf{x}(t) = T \circ \mathrm{f}(t) + \mathbf{n}(t) \tag{1.7}$$

$\mathbf{n}(t)$ is *white noise* if $R_{\mathbf{nn}}(\tau) = S_0\delta(\tau) \Leftrightarrow S_{\mathbf{nn}}(\omega) = S_0$. In general, this does not happen. However, under certain conditions[2], it is possible to represent the acquired noise signal as the output of a causal filter $L(t)$ fed with a white-noise process (or *innovations*) $\mathbf{i}(t)$ (see Figure 1). The following relations hold in general when $\mathbf{i}(t) = L \circ \mathbf{n}(t)$:

$$\mathrm{S}_{\mathbf{nn}}(\omega) = \mathrm{S}_{\mathbf{ii}}(\omega)|L(\omega)|^2 \quad \Leftrightarrow \quad \mathrm{R}_{\mathbf{ii}}(\tau) \equiv \int_{-\infty}^{+\infty} \mathrm{l}(t_1)\mathrm{l}(t_2)\mathrm{R}_{\mathbf{nn}}(\tau - t_1 + t_2) dt_1 dt_2 \tag{1.8}$$

The use of stochastic processes is somewhat unnatural when dealing with the sampled output of an analog-to-digital converter (ADC), even if the hypotheses of sampling theorem are satisfied and then the two descriptions are exactly equivalent. All our previous

[2] Paley-Wiener condition, see [11] p.402.

statements can be easily restated in terms of *discrete* stochastic processes, i.e. countable infinite set of RV $\mathbf{n}_i$, $\mathbf{x}_i$, etc. where the index *i* is a representation of the time in which the ADC acquired the *i*th sample $\mathbf{x}_i \equiv \mathbf{x}(t_i)$. From now hence, we shall admit that all signals we are dealing with are band-limited, and we have correctly dealt with the requirements on the minimum sampling frequency.

1.1.2 The Wiener-Kolmogorov filter

From the previous section, we know that the sampled signal after the ADC stage is $\mathbf{x}_i = T \circ \mathrm{f}_i + \mathbf{n}_i$, where $\mathbf{n}_i$ is a zero mean discrete stochastic process ($\langle \mathbf{n}_i \rangle = 0$). In a real experiment, we are given a finite length stream $\{\mathbf{x}_i\}_{i=1\ldots N}$, and we are required to develop a *linear* filter *W* capable of separating in some sense the signal f_i from the noise $\mathbf{n}_i$. To accomplish this task we shall put specific boundaries to our problem. Let us assume that:

A. The signal template $\{\mathrm{u}_i\}_{i=1\ldots N}$ is known[3], and the unique unknown is the *amplitude A* of the signal: $\forall i : \mathrm{f}_i = A\, \mathrm{u}_i$.

B. The autocovariance matrix of $\mathbf{n}_i$ (and hence of $\mathbf{x}_i$) is known:

$$\mathrm{R}_{ij} \equiv [\mathrm{R}_{\mathbf{nn}}]_{ij} = \langle \mathbf{n}_i \mathbf{n}_j \rangle = \langle (\mathbf{x}_i - \bar{\mathbf{x}}_i)(\mathbf{x}_j - \bar{\mathbf{x}}_i) \rangle = [\mathrm{R}_{\mathbf{xx}}]_{ij}. \tag{1.9}$$

Let $\mathbf{A} = \sum_j \mathrm{w}_j \mathbf{x}_j$ be the output of the filter. The two requirements to this estimate of the signal amplitude are:

C. The estimate is *unbiased*:

$$\langle \mathbf{A} \rangle = A \quad \Leftrightarrow \quad \sum_j \mathrm{w}_j \mathrm{u}_j = 1 \tag{1.10}$$

D. The variance $\sigma_{\mathbf{A}}^2$ of the estimate is at minimum

Combining the two requirements it happens that the signal-to-noise ratio (SNR) $|\mathbf{A}|/\sigma_{\mathbf{A}}$ is minimized as well. There is a unique well-known solution to this problem under the stated hypothesis, namely the *Wiener-Kolmogorov* (WK) filter. The coefficients of the filter can be are obtained for instance using the Lagrange multipliers method:

$$\begin{aligned}\exists \lambda \in \mathbb{R} : 0 &= \frac{1}{2}\frac{\partial \sigma_{\mathbf{A}}^2}{\partial \mathrm{w}_k} + \frac{\lambda}{2}\frac{\partial \langle \mathbf{A} \rangle}{\partial \mathrm{w}_k} = \langle (\mathbf{A} - A)\mathbf{x}_k \rangle + \frac{\lambda}{2}\langle \mathbf{x}_k \rangle = \sum_j \mathrm{w}_j \langle \mathbf{x}_j \mathbf{x}_k \rangle - A^2\, \mathrm{u}_k + \frac{\lambda}{2} A\, \mathrm{u}_k = \\ &= \sum_j \mathrm{w}_j \left(R_{jk} + \mathbf{A}^2\, \mathrm{u}_j \mathrm{u}_k \right) - A^2\, \mathrm{u}_k + \frac{\lambda}{2} A\, \mathrm{u}_k = \sum_j \mathrm{w}_j\, R_{jk} + \frac{\lambda}{2} A\, \mathrm{u}_k\end{aligned} \tag{1.11}$$

hence, after multiplying by $\mu_{ij} \equiv \mathrm{R}_{ij}^{-1}$ and deploying the boundary condition to get rid of λ, we get

$$\mathrm{w}_i = \frac{\sum_j \mu_{ij}\, \mathrm{u}_j}{\sum_{hk} \mu_{hk}\, \mathrm{u}_h \mathrm{u}_k} \tag{1.12}$$

The residual of $\mathbf{A}$ off the mean is

$$\mathbf{a}_i = \frac{\sum_j \mu_{ij}\, \mathrm{u}_i\, \mathbf{x}_j}{\sum_{hk} \mu_{hk}\, \mathrm{u}_h \mathrm{u}_k} - A = \frac{\sum_j \mu_{ij}\, \mathrm{u}_i\, \mathbf{n}_j}{\sum_{hk} \mu_{hk}\, \mathrm{u}_h \mathrm{u}_k} \tag{1.13}$$

[3] We assume that the samples are correctly shifted in time; in Sec. ... we shall extend the method for the case in which there is a residual uncertainty on the phase or on the time of arrival.

which is a zero-mean stochastic process with variance

$$\sigma_{\mathbf{a}}^2 = \left\langle \left[\frac{\sum_{ij} \mu_{ij}\, u_i\, \mathbf{n}_j}{\sum_{hk} \mu_{hk}\, u_h u_k} \right]^2 \right\rangle = \frac{1}{\left(\sum_{hk} \mu_{hk}\, u_h u_k\right)^2} \sum_{ijpq} \mu_{ij} \mu_{pq}\, u_i u_p \left\langle \mathbf{n}_j \mathbf{n}_q \right\rangle = \frac{1}{\sum_{hk} \mu_{hk}\, u_h u_k} \tag{1.14}$$

1.1.3 The Wiener filter as a linear system

Up to now the template was always assumed to be in phase with the signal hidden in the acquired data. Actually, in most occasions the *phase* or *time of arrival* of the signal is unknown, and whether a signal is present or not is something that has to be decided on the data themselves.

A straightforward procedure to get this information is to shift in time the template:

$$u_j(t_0) \equiv u(Tj - t_0) \quad \Leftrightarrow \quad u(t;t_0) \equiv u(t - t_0) \tag{1.15}$$

To each value of the time shift t_0 corresponds a different template, and therefore a different estimate for $\mathbf{A}$. From this point of view, the WK filter can be thought as a linear time-invariant system *W*, with the data stream x(t) as input and A(t) being its output:

$$\mathbf{A}(t) = \frac{\iint \mu(t'-t'')\, u(t'-t)\mathbf{x}(t'')dt'dt''}{\iint \mu(t'-t'')\, u(t'-t)u(t''-t)dt'dt''} = \frac{\iint \mu(t''-t')\, u(-t')\mathbf{x}(t-t'')dt'dt''}{\iint \mu(t'-t'')\, u(t')\, u(t'')dt'dt''} = W \circ \mathbf{x}(t) \tag{1.16}$$

hence

$$w(t) = \frac{\int \mu(t-t')\, u(-t')dt'}{\iint \mu(t'-t'')\, u(t')\, u(t'')dt'dt''} \quad \Leftrightarrow \quad W(\omega) = \frac{S^{-1}(\omega)\, \tilde{u}^*(\omega)}{\int S^{-1}(\omega)\left|\tilde{u}(\omega)\right|^2 d\omega} \tag{1.17}$$

In the frequency domain, the *W* filter is a sort of average of the different Fourier coefficients weighted by the reciprocal of the noise power spectral density. The transfer function $W(\omega)$ admits also an interesting decomposition as a cascade of two subsystems, a whitening filter *L* and the white noise version of the filter itself, the *mask M*. In fact, recalling that $S(\omega) \equiv S_{\mathbf{nn}}(\omega) = L(\omega)L^*(\omega)$, we can write

$$W(\omega) = L(\omega)\, M(\omega) \tag{1.18}$$

where

$$M(\omega) = \frac{L^*(\omega)\, \tilde{u}^*(\omega)}{\int \left|L(\omega)\, \tilde{u}(\omega)\right|^2 d\omega} = \frac{L \circ \tilde{u}^*(\omega)}{\left\|L \circ u\right\|^2} \quad \Leftrightarrow \quad m(t) = \frac{L \circ u(-t')}{\left\|L \circ u\right\|^2} \tag{1.19}$$

being the denominator just equal to $\sigma_{\mathbf{a}}^2$.

In other words, in the whitened data domain the WK filter is realized by simply convolving the input stream with the signal template: $\mathbf{A}(t) = \left(L \circ u(t)\right) * \left(L \circ \mathbf{x}\right)(t)$.

It is worth to notice that *L* is a causal filter, u(t) is a causal signal template, but *M* is a purely *anticausal* filter. For this reason it can be difficult to implement with analog devices, while it is easy with buffered digital signal processing computations.

Another interesting insight comes by observing that in presence of pure noise $\mathbf{a}(t) = W \circ \mathbf{n}(t)$, hence

$$R_{\mathbf{aa}}(\tau) = \left\langle W \circ \mathbf{n}(t') W \circ \mathbf{n}(t''+\tau) \right\rangle = \int R(\tau - t' - t'')\, w(t')\, w(t'')dt'dt'' \tag{1.20}$$

$$\Rightarrow S_{\mathbf{aa}}(\omega) = |M(\omega)|^2 = \sigma_{\mathbf{A}}^2 \frac{|\tilde{u}(\omega)|^2}{S(\omega)} \equiv \sigma_{\mathbf{A}}^2 \,\mathrm{snr}^2(\omega)$$

$\mathrm{snr}^2(\omega)$ is the squared SNR per unit frequency. It is an invariant function, because whichever linear filter is applied to both the signal and the noise, it cancels out in the their ratio. Equation (1.20) suggests that M is in fact a band-pass filter, that suppresses the Fourier components of the incoming data stream where $\mathrm{snr}^2(\omega)$ is low, therefore enhancing the SNR of the filtered signal $\mathbf{a}(t)$.

Once we have this new stochastic process $\mathbf{a}(t)$, we apply the definition of WK filter, and look for the value of the time $\mathbf{t}$ for which $\mathbf{a}(t)$ reach an extreme[4]. The two RV $\mathbf{t}$ and $\mathbf{a}(\mathbf{t})$ are the estimates of the time of arrival and amplitude of the signal, respectively. However, these are no more *linear* estimates, because it implies a maximum/minimum search. It can be shown however that it is approximately linear at least in the limit of high SNR (> 25).

Finally, a multiple signal presence in the data should be allowed for. If the signals are widely spaced in time, the extremes search should just become a *local* search. In the case two or more signals are packed closest than the duration of the signal template in the filtered data domain, $W \circ \mathrm{u}(t)$, than the correct procedure is to change the template from $\mathrm{u}(t)$ to $\mathrm{v}(t;\Delta t, r) \equiv \mathrm{u}(t) + r\,\mathrm{u}(t-\Delta t)$, thus adding the signal separation Δt and the amplitude ratio r as a new parameters.

To conclude this section, we shall give a formal definition for the bandwidth $\Delta\omega$ of a function $S(\omega)$:

$$\Delta\omega \equiv \frac{\left[\int S(\omega)d\omega\right]^2}{\int S^2(\omega)d\omega} \tag{1.21}$$

[4] We don't look just for the maximum because, as we shall see, $\mathrm{u}(t)$ is an oscillating wave pattern that can start moving in the alternate direction, that's why the sign of the amplitude is important, too. A small error in the time $\mathbf{t}$ could even reverse the sign of $\mathbf{a}(\mathbf{t})$, that's why we look for the maximum *absolute* value. It should be remarked that the WK filter guarantees that the variance of the noise is minimum, or that the SNR –*not the amplitude*– is maximum.

1.2 χ^2 test

In the hypothesis of strict Gaussian noise the WK filter is a maximum likelihood filter too. The case of a diagonal autocorrelation matrix can be generalized to a generic symmetric matrix. The logarithm of the likelihood happens to be distributed as a standard χ^2 variable, and can be used as a test statistic.

In presence of a signal in the data, if it is matched with the template, than the distribution of the log-likelihood is unchanged, otherwise the distribution is a non-centered χ^2. The dependence on SNR is at first order a quadratic power law.

1.2.1 Maximum likelihood and hypothesis testing

Let us return to the original issue of the 'best' choice of the estimator for the signal amplitude A, assuming all other parameters known. Another approach could have been the *maximum likelihood method* (MLM), which consists in finding the value $\hat{A}$, which maximizes the probability $P(x_1,\ldots,x_N;A)$ of having the sequence of samples $\{x_i\}_{i=1\ldots N}$ as the outcome of an experiment when the amplitude of the input signal was A.

For the sake of simplicity, we shall deal first with white noise. This is not a restricting assumption, as we can apply the whitening filter L to $\{\mathbf{x}_i\}_{i=1\ldots N}$ thus obtaining the new signal $\mathbf{y}_i = L\circ\mathbf{x}_i = L\circ G\circ \mathrm{f}_i + \mathbf{i}_i$. The expectation value of $\mathbf{y}_i$ is $L\circ G\circ \mathrm{f}_i = A\cdot L\circ \mathrm{u}_i$. As $\{\mathbf{y}_i\}_{i=1\ldots N}$ are a set of independent RV, and recalling that $\sigma_{\mathbf{y}}^2 = 1$, the likelihood function is simply

$$P(y_1,\ldots,y_N;A) \propto \exp\left[-\frac{1}{2}\sum\nolimits_i \left(y_i - A\,\mathrm{u}_i\right)^2\right] \qquad (1.22)$$

Taking the minimum of $-\log P(y_1,\ldots,y_N;A)$ as a function of A gives

$$0 = \sum\nolimits_i \mathrm{u}_i\left(y_i - \hat{A}\mathrm{u}_i\right) \quad\Rightarrow\quad \hat{A} = \frac{\sum_i \mathrm{u}_i\, y_i}{\sum_k \mathrm{u}_k^2} \qquad (1.23)$$

We revert now to $\{x_i\}_{i=1\ldots N}$, using the following relation for the discrete time version of $R_{\mathbf{ii}}(t)$:

$$[\mathrm{R}_{\mathbf{ii}}]_{ij} = \sum\nolimits_{hk}[\mathrm{R}_{\mathbf{nn}}]_{hk}\,\mathrm{L}_{hi}\,\mathrm{L}_{kj} \equiv \sum\nolimits_{hk}\mu_{hk}^{-1}\,\mathrm{L}_{hi}\,\mathrm{L}_{kj} = \delta_{ij} \qquad (1.24)$$

The likelihood function becomes

$$P(x_1,\ldots,x_N;A) \propto \exp\left[-\sum\nolimits_{ij}\mu_{ij}\left(x_i - \hat{A}\,\mathrm{u}_i\right)\left(x_j - \hat{A}\,\mathrm{u}_j\right)\right] \qquad (1.25)$$

and its maximum is found at

$$\hat{A} = \frac{\sum_{ij}\mu_{ij}\,\mathrm{u}_i\, x_j}{\sum_{hk}\mu_{hk}\,\mathrm{u}_h\mathrm{u}_k} \qquad (1.26)$$

It is not completely surprising that this estimate coincides with the output of the WK filter. The MLM is known to be[5] a minimum variance too, and we have proved that WK filter is the unique possible solution. This can be restated as:

The parameter that maximizes the probability of an event in the data represent the best estimate in terms of signal-to-noise ratio, as produced by a linear filter.

As the next step, we shall define a goodness-of-the-fit test of our estimate. As we know exactly the variance of the noise, and the statistics is Gaussian, then a χ^2-test fits our need. We shall therefore define in the whitened data the RV

$$\mathbf{X} = \frac{1}{n}\sum_i \left(\mathbf{y}_i - \mathbf{A}L^{-1}\circ \mathrm{u}_i\right)^2 \tag{1.27}$$

where n is the *number of degree of freedom*, in this case N-1.

As a function of the $\mathbf{x}_i$ variable, the expression of the test statistic $\mathbf{X}$ is

$$\mathbf{X} = \frac{1}{n}\sum_{ij}\mu_{ij}\left(\mathbf{x}_i - \mathbf{A}\,\mathrm{u}_i\right)\left(\mathbf{x}_j - \mathbf{A}\,\mathrm{u}_j\right) \tag{1.28}$$

Sometimes it can help also writing down the following equivalent expression:

$$\mathbf{X} = \frac{1}{n}\left[\sum_i \mathbf{y}_i^2 - \frac{\mathbf{A}^2}{\sigma_{\hat{A}}^2}\right] = \frac{1}{n}\left[\sum_{ij}\mu_{ij}\mathbf{x}_i\mathbf{x}_j - \frac{\mathbf{A}^2}{\sigma_{\hat{A}}^2}\right] \tag{1.29}$$

where the subtraction of the signal contribution has been singled out.

In presence of a signal matched with the template, this RV should obey a reduced χ^2 statistics:

$$f_{\mathbf{X}}(X;n,\sigma_{\mathbf{X}}^2) = \frac{1}{\left(2\sigma_{\mathbf{X}}^2/n\right)^{n/2}\Gamma(n/2)} X^{n/2-1} e^{-nX/2\sigma_{\mathbf{X}}^2} \tag{1.30}$$

This method modifies in a straightforward way in case the time of arrival of the signal, or even if other unknown parameters $\left\{\vartheta_j\right\}_{i=1\ldots P}$ enter (in principle non-linearly) in the signal template. However, no simple analytic formula provides us with the maximum of $P(x_1,\ldots,x_N;A,t_0,\left\{\vartheta_j\right\}_{i=1\ldots P})$, thus we should rely instead on numerical methods to maximize the likelihood function. It can be proved that χ^2-test statistic would be given in this case also by the logarithm of the likelihood function evaluated at the estimated values of the parameters:

$$\mathbf{X} = P\left(x_1,\ldots,x_N;\mathbf{A},\mathbf{t},\left\{\boldsymbol{\vartheta}_j\right\}_{i=1\ldots P}\right) \tag{1.31}$$

1.2.2 Power of the method

Let $\mathrm{u}'(t) \equiv L\circ \mathrm{u}(t)$ be the template after the whitening filter, and suppose that a different, *spurious* signal $\mathrm{v}'(t) \equiv L\circ \mathrm{v}(t)$ is present in the data. The event search procedure eventually produces an estimate of the amplitude of a gw event that is of course biased (it should have been zero!). We shall now derive the statistics of $\mathbf{X}$ in this case. In fact, if we are going to use $\mathbf{X}$ as a test statistic, then knowledge of the unbiased estimate statistics allow to compute the false dismissal probability, but a decision threshold usually require a balance between the

[5] In general, this holds in the large number of sample limit; but as we deal with Gaussian noise, it is an exact result.

latter and false alarm, which depends on the relative rate of the spurious signal and on the statistics of $\mathbf{X}$ when it is applied to one of them.

To proceed, let us make some remarks:

If the spurious signal has amplitude A, i.e. it is $A\mathrm{v}(t)$, then the mean value of $\mathbf{A}$ is

$$\langle \mathbf{A} \rangle = A \frac{\sum_{ij} \mu_{ij}\, \mathrm{u}_i\, \mathrm{v}_j}{\sum_{hk} \mu_{hk}\, \mathrm{u}_h \mathrm{u}_k} \equiv A\rho \tag{1.32}$$

- The variance of $\mathbf{a}_i$ is still given by (1.14), and it still true that $\sigma_{\mathbf{A}}^2 = \sigma_{\mathbf{a}}^2$, because we are assuming that there is just a spurious signal on the data, but with the same noise model.
- The residuals $\left(\mathbf{y}_i - \mathbf{A}\,\mathrm{u}_i'\right)$ are biased normal RV, while $\left(\mathbf{y}_i - \mathbf{A}\rho^{-1}\, \mathrm{v}_i'\right)$ are zero-mean normal RV, so their squared sum is a χ^2-distributed RV.

Therefore, starting from (1.27), we obtain

$$\begin{aligned} n\mathbf{X} &= \sum_i \left[\mathbf{y}_i - \mathbf{A}\rho^{-1}\, \mathrm{v}_i' + \mathbf{A}\left(\rho^{-1}\, \mathrm{v}_i' - \mathrm{u}_i'\right)\right]^2 = \\ &= \sum_i \left(\mathbf{y}_i - \mathbf{A}\rho^{-1}\, \mathrm{v}_i'\right)^2 + 2\sum_i \mathbf{A}\left(\mathbf{y}_i - \mathbf{A}\rho^{-1}\, \mathrm{v}_i'\right)\left(\rho^{-1}\, \mathrm{v}_i' - \mathrm{u}_i'\right) + \sum_i \left[\mathbf{A}\left(\rho^{-1}\, \mathrm{v}_i' - \mathrm{u}_i'\right)\right]^2 \equiv \\ &\equiv n\mathbf{X}' + 2\mathbf{C} + \mathbf{A}^2 \sum_{ij} \mu_{ij} \left(\rho^{-1}\, \mathrm{v}_i - \mathrm{u}_i\right)\left(\rho^{-1}\, \mathrm{v}_i - \mathrm{u}_i\right) \end{aligned} \tag{1.33}$$

where $\mathbf{C}$ is the term in the middle of the second line, and $\mathbf{X} = \frac{1}{n}\sum_i \left(\mathbf{y}_i - \mathbf{A}\rho^{-1}\, \mathrm{v}_i'\right)^2$. The latter follows a reduced χ^2 statistic, and whose mean do not depend on A. $\mathbf{C}$ is the product of two normal RV, one of which is not zero-mean. It can be split in the sum of a zero-mean normal RV and a residual *proportional* to a χ^2 RV. The last term follows a 1-degree-of-freedom χ^2 statistics.

More than the details of the statistic of this variable, it worth to notice that its mean value is $\langle \mathbf{X} \rangle = 1$ plus a term proportional to the square of the spurious signal amplitude:

$$\langle \mathbf{X} \rangle = 1 + \lambda \langle \mathbf{A} \rangle^2 \tag{1.34}$$

This means that, at least for high SNR, it is possible to remove effectively the spurious events detected in the filtered data. The numerical value of λ –hence how much ‘high’ the SNR should be– can be computed once a template for the spurious signals is given.

Of course, it is natural to ask what happens to a *true* signal and its test statistic when the wrong filter has been applied, either because the noise model is wrong, or the parameters for the transfer function of the detector are misestimated. This should not happen! In the next Chapter we are going to discuss the methods by which the compliance of the WK filter with the noise PSD are tested. Moreover, the transfer function parameters are known to be a subset of the noise model parameters. Anyway, a few times a calibration signal was injected in the system to verify this assumption (see 4.1.6).

1.3 Event search without a trigger

For a multiple mechanical oscillator model, the pulse response after the WK filter shows a characteristic dumped oscillating pattern with beat modulation

A list of candidate events is built by looking for every independent local extreme of the amplitude modulation. This task is conveniently performed by a max-hold algorithm, with a characteristic time of the order of the Wiener time. To estimate accurately the amplitude and time of arrival A e t_0 *the WK filtered data have to be interpolated. Expected amplitude distribution of candidate events in absence of signal is discussed.*

Even neglecting interpolation errors, there are unavoidable peak and phase errors in the time of arrival estimate. The first ones depend on WK filter characteristic time or on the beat time between the resonating modes for small-bandwidth detectors. Phase errors depend on the carrier frequency. Approximate SNR scaling laws for peak and phase errors are derived.

1.3.1 A parametric model for coupled mechanical oscillators

A device that can be modeled as a linear electric network of passive elements, each one with simple (algebraic) transfer function, has an overall impulse response whose Laplace transform is a rational algebraic function. Its poles are not more than the number of topologically independent loops in the network, and their position depends only on the choice of the output port, and not on the placement of the external driving signal or internal noise sources. We shall give a brief demonstration of this statement.

No matter how complex the finite element model of a real *passive linear* system, in the frequency domain it is described as a system of algebraic equations, which can be solved with a matrix inversion to give the measurable quantities, i.e. output current or voltages. The typical steps in this procedure are the following:

1. write down the measured output voltage V_{out} as a linear combination of a specified set of independent loop currents $(i_1,\cdots,i_N)=\mathbf{I}$ and loop impedances $(z_1,\cdots,z_N)=\mathbf{Z}$ derived from the equivalent electric model of the system:
$$V_{out}=\mathbf{I}\cdot\mathbf{Z} \tag{1.35}$$
2. write the dynamic equations for each loop, and solve the circuit for the currents *I*, as a function of the voltage sources on each loop $(v_1,\cdots,v_N)=\mathbf{V}$:
$$\mathbf{M}\cdot\mathbf{I}=\mathbf{V} \quad\Rightarrow\quad \mathbf{I}=\frac{1}{\Delta}\mathbf{N}\cdot\mathbf{V} \qquad\qquad \Delta\equiv\det\mathbf{M},\quad \frac{\mathbf{N}}{\Delta}=\mathbf{M}^{-1} \tag{1.36}$$
3. express the output voltage V_{out} as a function of the voltage sources in the network:
$$V_{out}=\frac{1}{\Delta}\sum_{ij}Z_iN_{ij}V_j=\frac{Q(V)}{P} \tag{1.37}$$

Note that if $p_k\in\mathbb{C}$ is a zero of the denominator *P*, then p^* is also a zero, as the inverse Laplace transform of all these functions is supposed to be a real function.

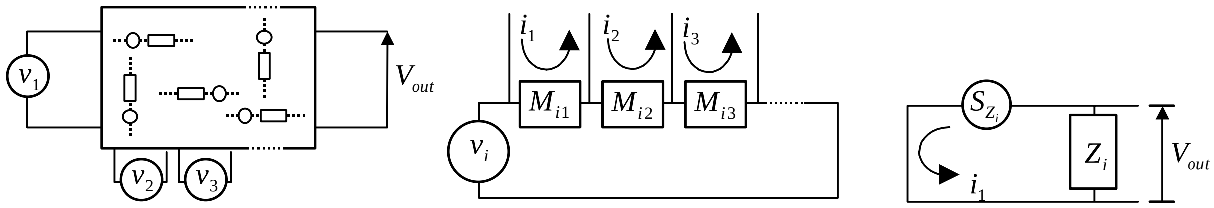


Figure 2 – (*left*) A linear physical device –like a bar detector– can be modeled as a network of linear passive elements, driven by voltage sources either deterministic or stochastic. (*center*) All the currents circulating inside the network can be described in terms of a set of independent loop currents i_k driven by the sources v_h through the impedances M_{hk} and by the Nyquist noise sources (*right*) linked to the output coupling impedance Z_h of each loop.

A similar result holds also for the PSD of the noise at the output. The Nyquist theorem relates the noise sources of the system to the dissipative (i.e. real) part of the output impedances $\{Z_n\}$ through the temperature T and the Boltzmann constant k_B:

$$S_{Z_i}(\omega) = 2k_B T \,\mathrm{Re}\{Z_i\} \tag{1.38}$$

They act as random voltages $\{V_n\}$ added to the signals. At the output port of the system, the spectral density due to each single noise source superimpose, weighted by the square of the transfer function of the signals $\{V_n\}$ (see (1.6)):

$$S_{n_{out}}(\omega) = \frac{1}{|\Delta|^2} \sum \left| Z_i N_{ij} \right|^2 S_{n_j}(\omega) \tag{1.39}$$

Note that each zero p of $|\Delta|^2$ in the negative real component half-plane comes with the corresponding positive real component counterparts $-p_k$: this is a standard feature of a PSD, as it is defined as the FT of an even function, see (1.4) and (1.5). In the following, whenever we refer to *the* pole p, it will imply also p^*, $-p$, $-p^*$.

From (1.37) and (1.39) we see that –provided that the transfer function of the basic elements of the system are rational polynomial of the complex frequency– the analytical models for the both the transfer function and the noise PSD of the whole system are always ratio of polynomials. What is more, the poles of the transfer function are a subset of the poles of the PSD.

Each of the poles p corresponds *by definition* to a *resonance* of the system. In the usual notation $p = i\omega_0 - \Delta$, $\omega_0/2\pi$ is the frequency of the resonance, Δ/π is the bandwidth (defined as full-width-half-maximum of the peak in the *energy* spectrum). Due to typically small bandwidth relative to frequency, $\Delta \ll \omega$, the resonance frequencies are usually well separate, and can be studied individually in a small frequency band.

As it was briefly explained in the introduction, bar detectors now in operation consist of a resonant mass (the bar) coupled to an electromechanical resonant transducer with the same free resonance frequency. The output of the detector –after a linear (or linearized) amplification stage– is acquired with an ADC, either directly (as in AURIGA) or through a band-pass amplifier. To our purposes, two coupled damped harmonic oscillators can conveniently model the entire system. The solution of the dynamical equations for the normal modes is a doublet of well-separated resonance modes (at least for present working detectors setup, ~20Hz at ~1kHz).

Close to the kth resonance frequency $\omega_k (k=1,2)$, the transfer function of the system resembles that of a harmonic oscillator, namely $\left(\omega^2 - \frac{i}{Q_k}\omega_k\omega - \omega_k^2\right)^{-1} \equiv \left[\left(i\omega - p_k\right)\left(i\omega - p_k^*\right)\right]$, where Q_k is the mechanical quality factor of the kth mode and the pole $p_k \equiv i\omega_k - \Delta_k$ ($\Delta_k > 0$)

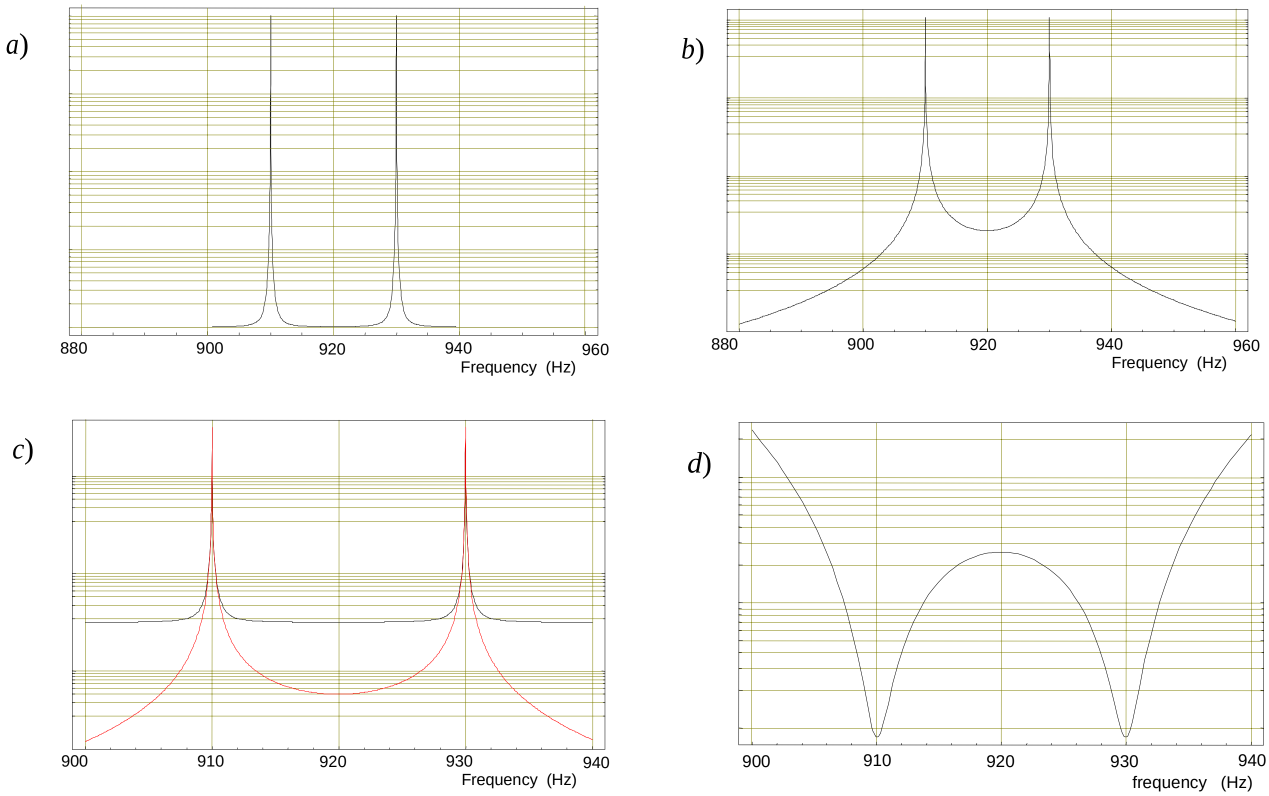


Figure 3 – The noise model for a 2-modes resonant detector similar in performance to AURIGA is shown in (*a*), while (*b*) represent the squared transfer function for the system in energy. Both functions are characterized by the same mechanical quality factor; therefore the real bandwidth of the system is much more than the inverse of the relaxation time of the mechanical oscillators (~1000 s). (*c*) The sensitive 'bandwidth' is the width of the lorenzian peak at the level of the amplifier wide band noise, up to which the two functions superimpose almost perfectly. This shows up promptly in (*d*), where the ratio of (*a*) over (*b*) is plotted. This function, once properly calibrated by a constant factor, gives the modelled noise at the input of the detector.

is placed in the negative real component half-plane of the complex plane, because the system has a causal impulse response. Hence we can model the total transfer function as

$$G(\omega) = \alpha \prod_{k=1}^{2} \frac{i\omega}{(i\omega - p_k)(i\omega - p_k^*)} \tag{1.40}$$

where α is a calibration constant to be experimentally determined.

Each dissipative element has an associated the *thermal noise* source through the Nyquist theorem, whose output power spectral density (PSD) is proportional to

$$\left|\omega^2 - \tfrac{i}{Q_k}\omega_k\omega - \omega_k^2\right|^{-2} \equiv \frac{1}{(i\omega - p_k)(i\omega + p_k)(i\omega - p_k^*)(i\omega + p_k^*)} \tag{1.41}$$

This is a *Lorentzian spectrum*, and it can be used in general to fit the PSD in a frequency range close to the resonances of the system.

Along with thermal noise, we should take care of the back action and of the wide band noise of the amplifier. In the context of this chapter, with the detector being modeled by serial electrical RLC oscillators read by a voltage amplifier, the *back action* of the amplifier itself is modeled by a current noise source acting from the output port of the system. As the output admittance of the system shares the same poles p_k of the transfer function, the voltage noise induced by the back-action acts just like an additive contribution in excess to thermal noise. The total output noise PSD $S_{nn}(\omega)$ is made by a linear combination of Lorentzian terms, one for each resonance, plus a constant *wide band noise* S_0 due to the voltage noise of the electronic amplification stage:

$$S_{\mathbf{nn}}(\omega) = S_0 + \sum_{k=1}^{2} \frac{C_k}{(i\omega - p_k)(i\omega + p_k)(i\omega - p_k^*)(i\omega + p_k^*)} \tag{1.42}$$

for suitable C_k. Arranging all terms, and using symmetry properties of the noise PSD, we conclude that the general expression for $S_{\mathbf{nn}}(\omega)$ is

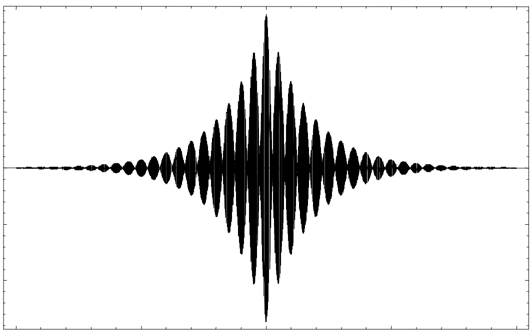
$$S_{\mathbf{nn}}(\omega) = S_0 \prod_{k=1}^{2} \frac{(i\omega - q_k)(i\omega + q_k)(i\omega - q_k^*)(i\omega + q_k^*)}{(i\omega - p_k)(i\omega + p_k)(i\omega - p_k^*)(i\omega + p_k^*)} \tag{1.43}$$

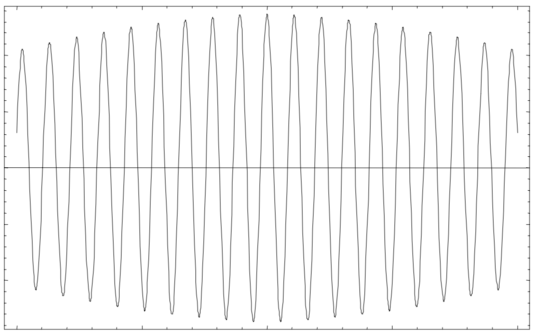
The complex zeros can be written as $q_k \equiv i\omega_k - \delta_k$ where $\mathrm{Re}\, q_k = \delta_k/\pi$ is the effective post-filtering bandwidth, and $\mathrm{Im}\, q_k \approx \mathrm{Im}\, p_k \equiv \omega_k$ if the correlation between wide-band noise and narrow-band noise is negligible (in particular, if $\Delta_k \ll \delta_k$). It is remarkable that these are all the parameters we need in order to build the WK filter for an impulsive signal, compared to the thousands (or more) involved with a *non-parametric* estimate of the power spectrum. On the other side, this approach gives consistent results only when the real noise PSD of the detector reasonably follows the model (e.g. it doesn't exhibits extra noise resonances about the normal modes).

Once the model is established, it is straightforward to build up the optimal WK filter for it[6], and all we need is to specify the signal template. Unless stated differently, in the following we shall focus to the response $u_\delta(t)$ to a δ-like GW signal entering the system (in a similar fashion we characterize a linear system with its impulse response). The FT $\tilde{u}_\delta(\omega)$ of the signal is simply the total transfer function

$$\tilde{u}_\delta(\omega) = \alpha \prod_k \frac{i\omega}{(i\omega - p_k)(i\omega - p_k^*)} \quad . \tag{1.44}$$

Figure 4. In the time domain, the signal after the WK filter appears as a damped beat between the two modes modulating the sinusoidal carrier wave. With a bandwidth of (~1Hz) typical of the present class of operating detectors, the relative amplitude of the peaks near the time of arrival is decreasing slowly with a quadratic low, because the beat modulation is faster than the exponential decay.





It's worth to remark that the response of the system to any other input signal h*(t)* can be always written as the time convolution $u_\delta * h(t)$. Therefore, once applied the optimum *linear* filter matched with u_δ, we can do WK filtering matched to *h* simply by convolution of h(*t*) with the data filtered for the impulse response: $(h * u_\delta) * \mathbf{x} = h * (u_\delta * \mathbf{x})$.

---

[6] Though it is handy to describe the effect of the WK filter in the frequency domain, yet the real implementation is better done in the discrete time domain with an equivalent autoregressive and moving average (ARMA) type filter implementation, see appendix.

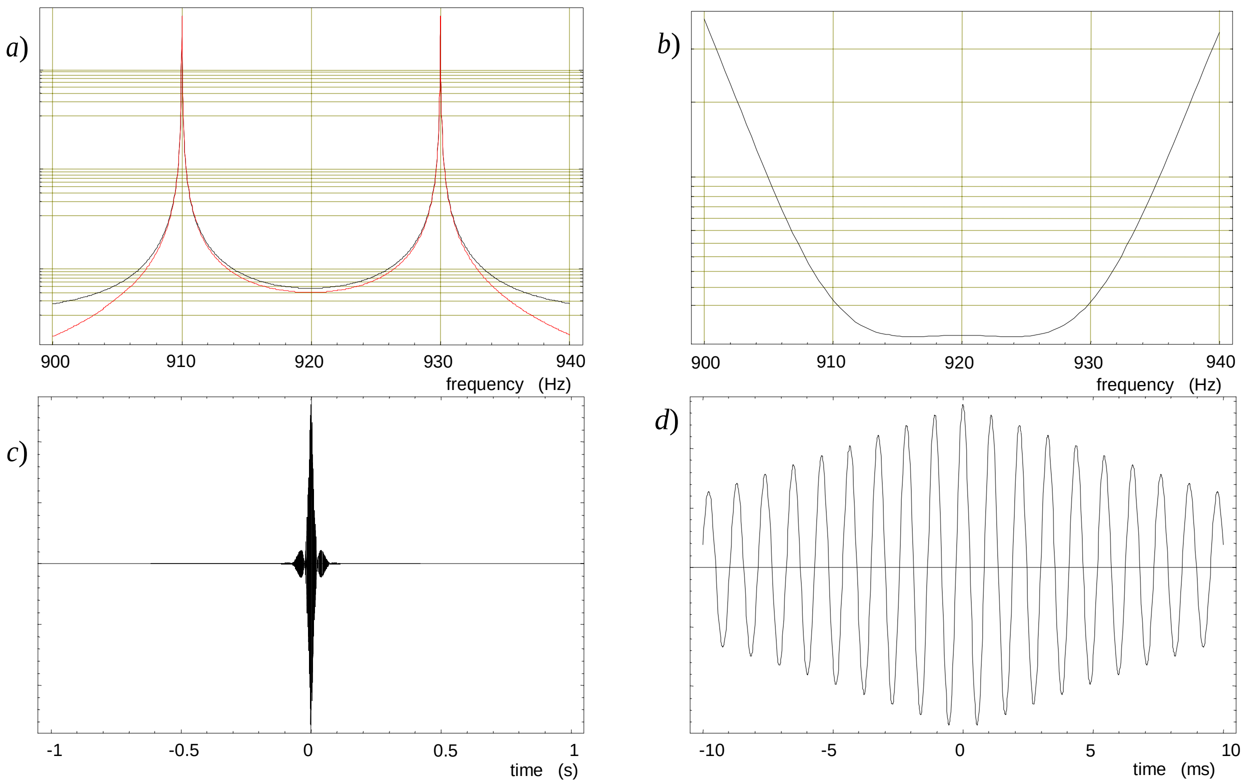


Figure 5 – In this example, a ~10Hz bandwidth per mode detector is examined, which is likely to be the conservative minimum achievement of the forthcoming upgraded bar detectors. In (*a*) the modelled noise is superimposed to the squared transfer function, and the equivalent gw noise spectral density is shown in (*b*). The two mode sensitive bands are no more separate, as they merge in a wide, flat bottom in the spectrum. The timescale of the beating is now the same or less of the exponential decay. As a consequence the side peaks near the time axis origin decay much faster, improving by a large factor the timing accuracy (see 1.3.3).

A more physical reason to focus on δ-like signals is that astrophysical models of impulsive GW sources suggest that duration of the signal component at 1kHz is of the order of ~1 ms. When such short burst are detected with the comparatively small bandwidth (~1÷50Hz) of resonant detectors, the exact details of the time structure are lost.

The expression of the WK filter in the frequency domain (1.18) specialize in

$$\mathrm{W}(\omega)=\sigma_{\mathbf{a}}^2 S_0^{-1}\prod_{k=1}^{2}\frac{-i\omega\left(i\omega-p_k\right)\left(i\omega-p_k^*\right)}{\left(i\omega-q_k\right)\left(i\omega+q_k\right)\left(i\omega-q_k^*\right)\left(i\omega+q_k^*\right)}. \qquad (1.45)$$

It is rather obvious how to split up the WK filter in the whitening filter (i.e. $|L(\omega)|^2\,\mathrm{S}_{\mathbf{nn}}(\omega)=1$)

$$\mathrm{L}(\omega)=S_0^{-1/2}\prod_{k=1}^{2}\frac{\left(i\omega-p_k\right)\left(i\omega-p_k^*\right)}{\left(i\omega-q_k\right)\left(i\omega-q_k^*\right)} \qquad (1.46)$$

and the mask:

$$\mathrm{M}(\omega)=\sigma_{\mathbf{a}}^2 S_0^{-1/2}\prod_{k=1}^{2}\frac{-i\omega}{\left(i\omega+q_k\right)\left(i\omega+q_k^*\right)} \qquad (1.47)$$

The latter is clearly a band-pass filter around the frequencies $\{\omega_k\}_{k=1,2}$.

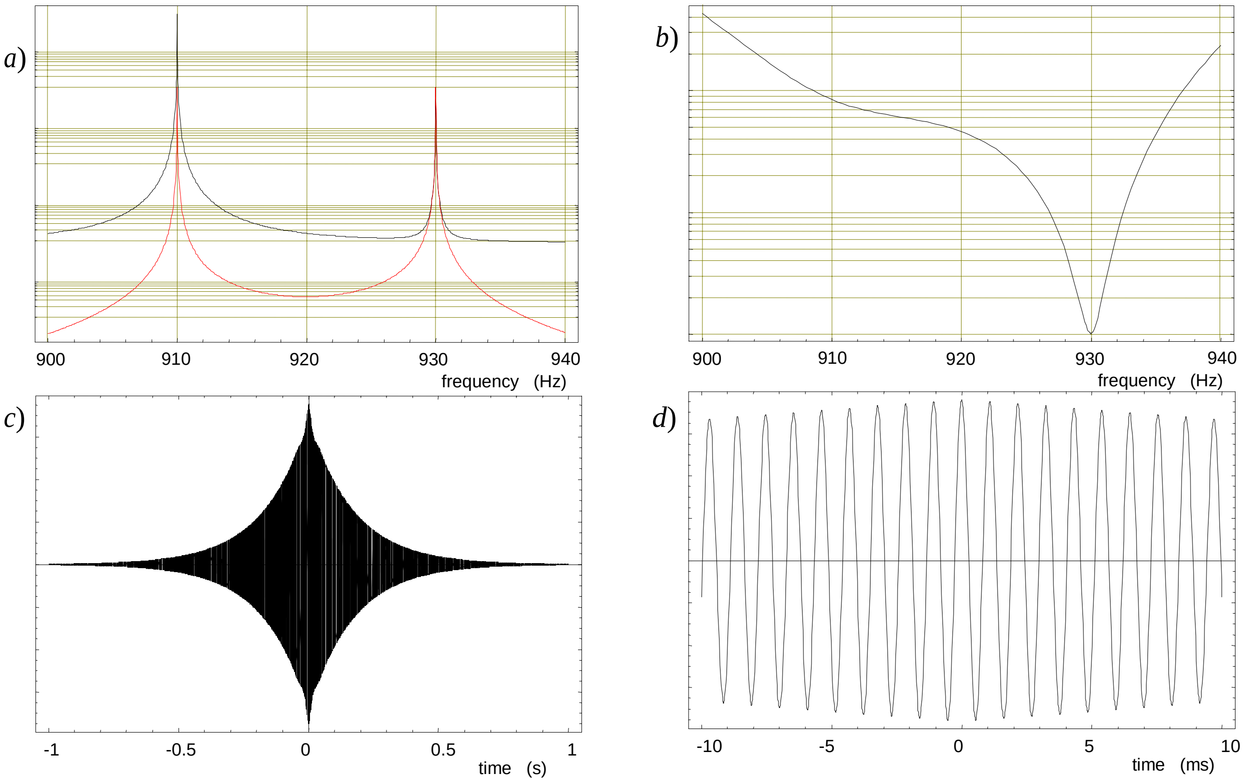


Figure 6 – (*a*) The two modes of a real detector are known to behave quite differently with respect of noise properties, as they interact with the other environmental disturbances (see Figure 60). We consider here a pessimistic case of a factor 10 in the relative post filtering bandwidth parameter –i.e. $\Re(q_1) \approx 10\Re(q_2)$. (b) The resultant noise at the input of the detector basically shows only one mode is sensitive to gw in comparison with the case of Figure 4. (*c,d*) The timing accuracy is slightly (a factor two) worse at high SNR, when only a few side peaks can play with the noise. At lower SNR, without the intervening beating, the exponential decay is much slower.

When a GW *burst* with amplitude A impinges on the bar, the FT of the signal after WK filter is

$$W \circ \tilde{u}_\delta(\omega) = A\sigma_{\mathbf{a}}^2 S_0^{-1} \prod_{k=1}^{2} \frac{\omega^2}{(i\omega - q_k)(i\omega + q_k)(i\omega - q_k^*)(i\omega + q_k^*)} \tag{1.48}$$

Comparing (1.44) and (1.48) we see that the WK filter changes the *pre-detection bandwidth* $\Delta\omega_k \equiv -\mathrm{Re}(p_k)$ of the modes to the *post-filter bandwidth* $\Delta\omega_k^{opt} \equiv -\mathrm{Re}(q_k)$. Moreover, as was remarked in 1.1.3, the two poles $-q_k$ and $-q_k^*$ introduce an noncausal component in the response.

1.3.2 Event search

We shall now revert to the time domain the results achieved in the previous section. From (1.48), it can be seen that a candidate δ-like event in the WK filter output is[7] a pattern with a specific mix of an exponential decay *both forward and backward* in time with time constant $t_W \approx 1 \big/ \max_{k=1,2}\{\Delta\omega_k^{opt}\}$ (*Wiener time*), the *beat* between the two normal modes at the frequency $\omega_* \approx \frac{1}{2}[\omega_2 - \omega_1]$ and a carrier wave at the frequency $\omega_0 \approx \frac{1}{2}[\omega_1 + \omega_2]$:

$$W \circ u_\delta(t) \approx Ae^{-t/t_W} \cdot \cos(\omega_* t)\cos(\omega_0 t) \tag{1.49}$$

[7] Actually, the exact solution is a superposition of two decay patterns with different decay times. But if they are almost equal then (1.49) is a useful approximation.

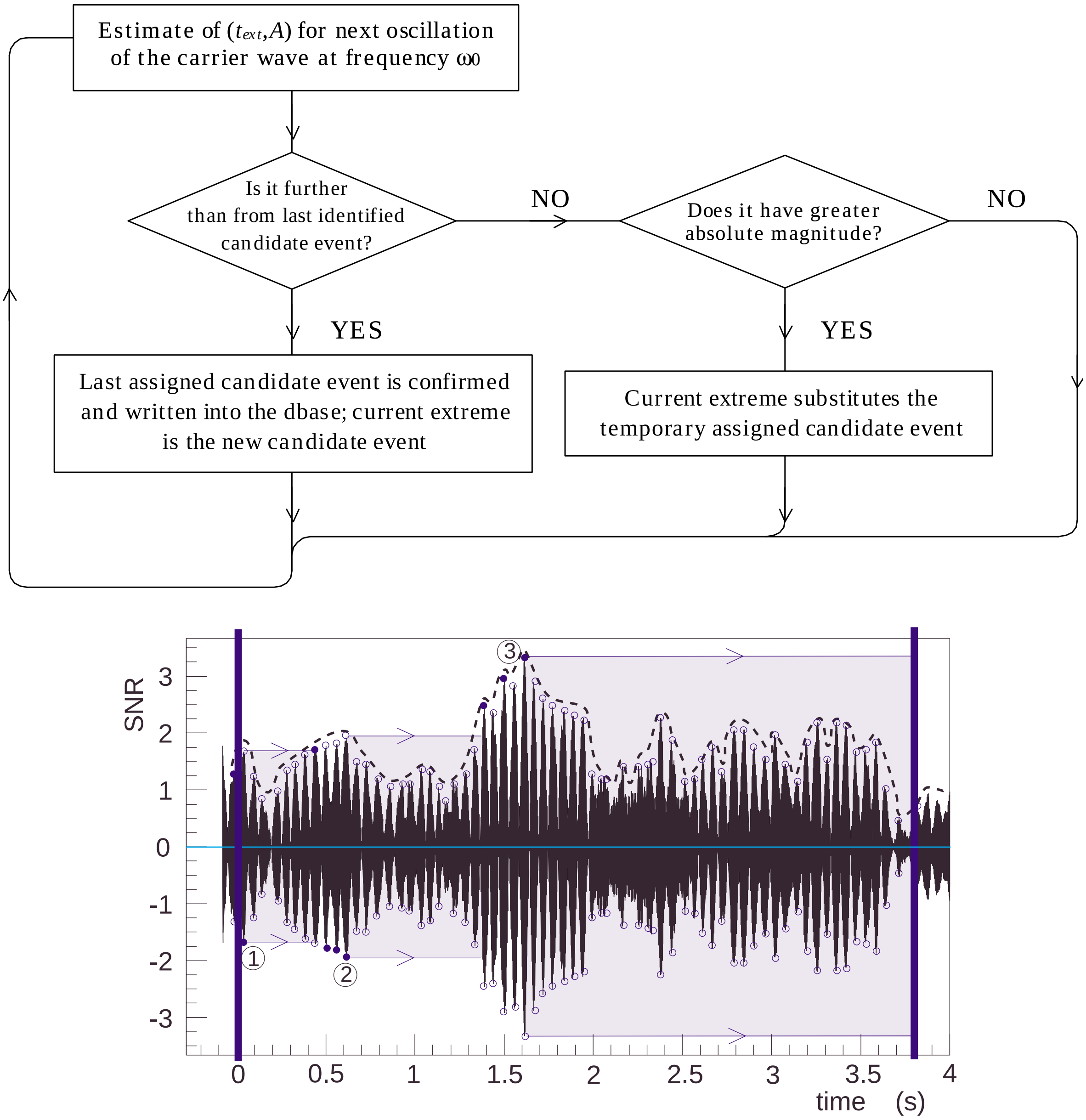


Figure 7. Flow chart of the event search algorithm (*top*) and corresponding pictorial representation (*above*), using a filtered stretch of noise with the same parameters of Figure 3. The extreme of a single beat cycle with amplitude exceeding that of all previous ones is marked by a solid filled dot, the others by an empty one (the individual extremes of the carrier wave are not visible on this timescale). The search for an event is renewed $\sim 3t_w$ past the last biggest event ③ –which we call then a candidate event–, while in ① and ② the extremes were not hold enough time. It is worth to remark that the event found in this way is not a real "signal", but just a statistical fluctuation of the background stochastic process. There are no explicit amplitude thresholds applied in the event search algorithm.

It is obvious that the local extremes of the filtered signal shown in Figure 4 are the minima and maxims of the carrier wave, whose coordinates (t_{ext},A) can be found accurately by interpolating the signal beyond the Nyquist sampling time, as we describe extensively in 4.1.3. However, only a subset of these extremes is *uncorrelated*, and we have to take care that we do not count the same event more than once.

A *maximum-hold algorithm* (see Figure 7) performs a search for local *uncorrelated* extremes, selecting only one of them in a suited interval of time. The Wiener time seems a good choice, as represents the decay time of the filtered signal (and also the noise autocorrelation time), and it fixes a lower bound on the separability of two temporally close pulses. In other words, because of the correlation in the data after the WK filter, we are forced to introduce a "dead time" $\sim t_W$ around each event, because we have to wait that the signal has decayed into the noise before looking for next event. We can compute the time needed by the template $W \circ \mathrm{u}_\delta(t)$ to decay below SNR=1 using the formula $\tau = t_W \ln SNR$. For instance, a signal with SNR=20 at the maximum needs at least $3t_W$ for its amplitude to decrease under SNR=1.

Obviously, when a signal has an intrinsic finite duration longer than the decay time of the WK filter it is erroneously classified as a sequence of impulsive events (but they are readily rejected as spurious by the χ^2 test).

We should be aware that the event search algorithm described here is going to produce a lot of false alarms at low SNR, because it recognizes as a candidate gw signal every single fluctuation of the stochastic process $W \circ \mathbf{n}(t)$ in a time span of duration τ. This is however unavoidable with no high threshold set, and justifies us calling gw detectors "noise dominated", meaning that most of the background comes from a continuous stochastic process, while apparently similar devices like particle detectors are most of the time signal dominated, i.e. most of the background is due to spurious detection due to transient excitations, for instance because of cosmic ray showers.

It would be interesting to have a prediction for the amplitude density function of this bulk of background events. It should be a functional of the autocorrelation function, the SNR and of the decorrelation threshold (which in our case is again proportional to the WK filter bandwidth).

Just for sake of discussion, we shall now try to answer directly to the question.

Let us start with an even subdivision $\{t_k | k = 0,...,2N\}$ of the time interval $[t_0 - \frac{1}{2}\tau; t_0 - \frac{1}{2}\tau]$ in 2N parts. Take a normal stochastic process $\mathbf{x}(t)$ satisfying $\langle \mathbf{x}(t) \rangle = 0$ and $\langle \mathbf{x}^2(t) \rangle = 1$, with autocovariance function $\mathrm{R}(\tau)$. Each of the 2N+1 samples $\mathbf{x}_k \equiv \mathbf{x}(t_k)$ is a normal RV, and the generalized (2N+1)-dimensional density function for the multivariate RV $\mathbf{x} \equiv (\mathbf{x}_{-N},...,\mathbf{x}_N)$ is

$$f_{\mathbf{x}}(x_{-N},...,x_N) = \frac{1}{\sqrt{(2\pi)^{2N+1}\Delta}} \exp\left(-\frac{1}{2}\mathbf{x}^T \boldsymbol{\mu} \mathbf{x}\right) \tag{1.50}$$

where

$$\boldsymbol{\mu} \equiv R^{-1}; \qquad R \equiv \{R_{hk} \equiv R(t_h - t_k)\}; \qquad \Delta \equiv \det \mathbf{R}$$

The probability that x_0 itself is a local maximum when it assumes the value a is

$$\begin{aligned} P(x_0 > x_k | x_0 = a) &= \iiint_{|x_k|<a} \delta(x_0 - a) f_{\mathbf{x}}(x_{-N},...,x_N) = \\ &= \iiint_{|x_k|<a} dx_{-N}...dx_N \delta(x_0 - a) \frac{e^{-\frac{1}{2}\mathbf{x}^T \boldsymbol{\mu} \mathbf{x}}}{\sqrt{(2\pi)^{2N+1}\Delta}} \end{aligned} \tag{1.51}$$

It is convenient to express this formula having the integrand depending on $\mathbf{R}$ (which has better symmetry properties than $\boldsymbol{\mu}$); this can be done changing the variable $\mathbf{x}$ into the Fourier transform conjugate variable $\Omega = (\omega_{-N},...,\omega_N)$:

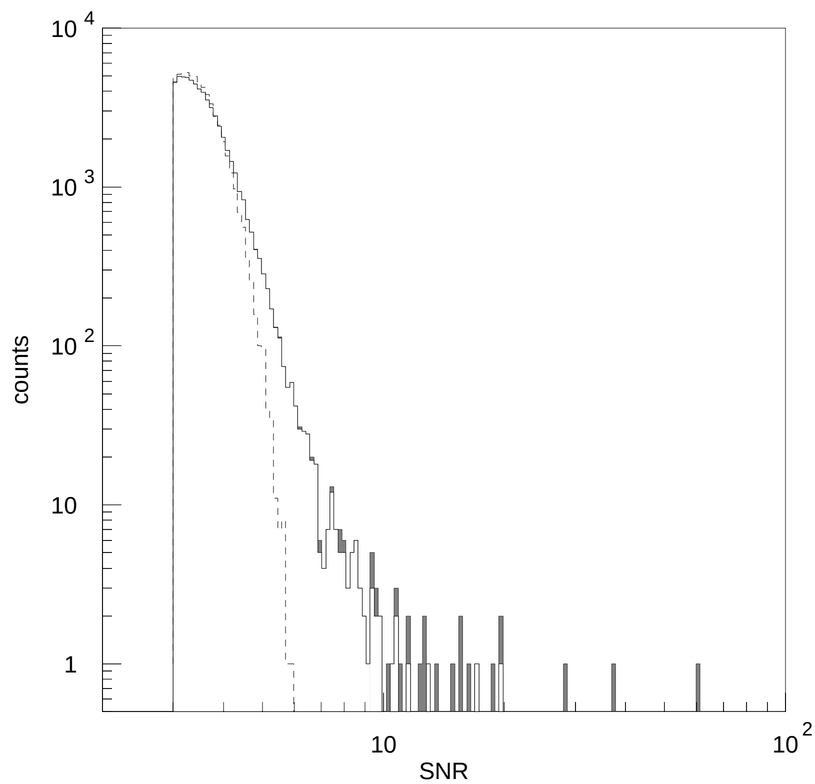
$$
\begin{aligned}
P(x_0 > x_k \middle| x_0 < a) &= \iiint_{x_k \in \mathbb{R}} dx_{-N} \dots dx_N \frac{e^{-\frac{1}{2}\mathbf{x}^T \boldsymbol{\mu} \mathbf{x}}}{\sqrt{(2\pi)^{2N+1}\Delta}} \delta(x_0 - a) \prod_{k \neq 0} \left[\vartheta(x_k + a) - \vartheta(x_k - a) \right] = \\
&= \iiint_{\omega_k, x_k \in \mathbb{R}} d\omega_{-N} \dots d\omega_N dx_{-N} \dots dx_N \frac{e^{-\frac{1}{2}\mathbf{x}^T \boldsymbol{\mu} \mathbf{x}}}{\sqrt{(2\pi)^{2N+1}\Delta}} e^{i(x_0 - a)\omega_0} \prod_{k \neq 0} \frac{1}{\omega_k} \left[e^{i(x_k + a)\omega_k} - e^{i(x_k - a)\omega_k} \right] = \\
&= \iiint_{\omega_k \in \mathbb{R}} d\omega_{-N} \dots d\omega_N e^{-\frac{1}{2}\Omega^T \mathbf{R} \Omega} e^{-ia\omega_0} \prod_{k \neq 0} \frac{2\sin(a\omega_k)}{\omega_k} = \\
&= \iiint_{\xi_k \in \mathbb{R}} d\xi_{-N} \dots d\xi_N e^{-\frac{s^2}{2}\Omega^T \frac{\mathbf{R}}{R_{00}} \Omega} e^{-i\xi_0} \prod_{k \neq 0} \frac{2\sin \xi_k}{\xi_k}
\end{aligned}
\qquad (1.52)
$$

where $s \equiv a \cdot \sqrt{R_{00}}$ is the SNR of the signal with amplitude *a*. The solution to the last integral in the limit $N \to \infty$ is the density function we are looking for.

Unfortunately, it seems we have found an integral that has no known solutions. And it is difficult to cope with it even numerically!

There is a cheap alternative to obtain the same result, using a Monte Carlo method. This means that we have to *simulate* the output of a WK filter with the same PSD, for instance by feeding an appropriate linear system with a Gaussian white noise generator (this can be done efficiently with an ARMA implementation, see 4.1.4). Then using a fast off-line event search algorithm we get a large number of events, whose histogram gives a fair estimate of the density function we are looking for. The result is shown in

Figure 8 – (*continous line*) Histogram of event amplitudes for 10 days of AURIGA data (UTC 12÷21 Jun 1997), and (*dashed line*) the analogous when the analysis is fed by a simulated Gaussian noise. To have a fair comparison, the parameters of the simulated noise are non-stationary, and use the parameters estimated hour by hour for the corresponding time span in AURIGA data.



1.3.3 Timing errors

Looking at Figure 4, it is rather easy to figure out what is the qualitative behavior of the timing error. First of all, if the signal has very low amplitude, then the noise alters the relative amplitude of the single beatings, and the central lobe actually can be found at lower amplitude than the side lobes, through this case is partially suppressed by the exponential decay, which of course favors the central lobe. As the SNR increases, the step needed for the noise to perform a lobe flipping is greater, until the chances are so low that basically the found time of arrival stick to the central lobe.

The same reasoning can be applied in a finer scale to the single oscillations of the carrier wave. Again, if the SNR is low, the noise is able to promote as local maximum any of the extremes near the top of the central beating lobe.

Eventually, at very high SNR, even amplitude flipping between nearby peaks is forbidden, and the timing error becomes a fraction of the period $T_0 = \omega_0^{-1}$ of the carrier wave.

We can summarize what we said writing the timing error as

$$\mathbf{\Delta t} = \mathbf{\Delta t_\phi} + \mathbf{k} T_* \tag{1.53}$$

where $\mathbf{\Delta t}$ is the total timing error, which is always (by definition!) less than the decorrelation time of the event search algorithm (which in turn is a multiple of the renewal time, i.e. the Wiener time t_W); we call $\mathbf{\Delta t_\phi}$ and $\mathbf{k}$ respectively the *phase error*; and the *peak error* (or *peak number*). They are defined as the unique solution of (1.53) such that $T_0/2 < \mathbf{\Delta t_\phi} < T_0/2$ and $\mathbf{k} \in \mathbb{Z}$.

The problem of finding the joint density function of $\mathbf{\Delta t_\phi}$ and $\mathbf{k}$ is not less involved than the exact solution we were looking for in the previous paragraph for the sole amplitude distribution. However, it turns out that if we are just interested to a solution for large SNR (i.e. small timing errors) then it can be proved [10] that there is a linear approximation to the problem. The solution is the following:

- $\mathbf{\Delta t_\phi}$ is a normal variable whose standard deviation does not depend on $\mathbf{k}$:

$$\sigma_\phi = \frac{1}{\omega_0 SNR} \tag{1.54}$$

  which evaluates to $\sigma_\phi \approx 173\mu s / SNR$ for present AURIGA hardware setup.

- The probability of selecting the maximum $\mathbf{k}$ depends from the product $t_W \cdot \omega_*$:

$$\begin{aligned} t_W \cdot \omega_* \ll 1 \quad &\Rightarrow \quad \sigma_k = \frac{t_W \cdot \omega_0}{SNR^2} \\ t_W \cdot \omega_* \gg 1 \quad &\Rightarrow \quad \sigma_k = \frac{\omega_0}{\pi \omega_* SNR} \end{aligned} \tag{1.55}$$

These two cases are illustrated by Figure 4, where the exponential decay is slower than the beating time, and by Figure 5, where the opposite case take place.

2 Adaptive models and the real world

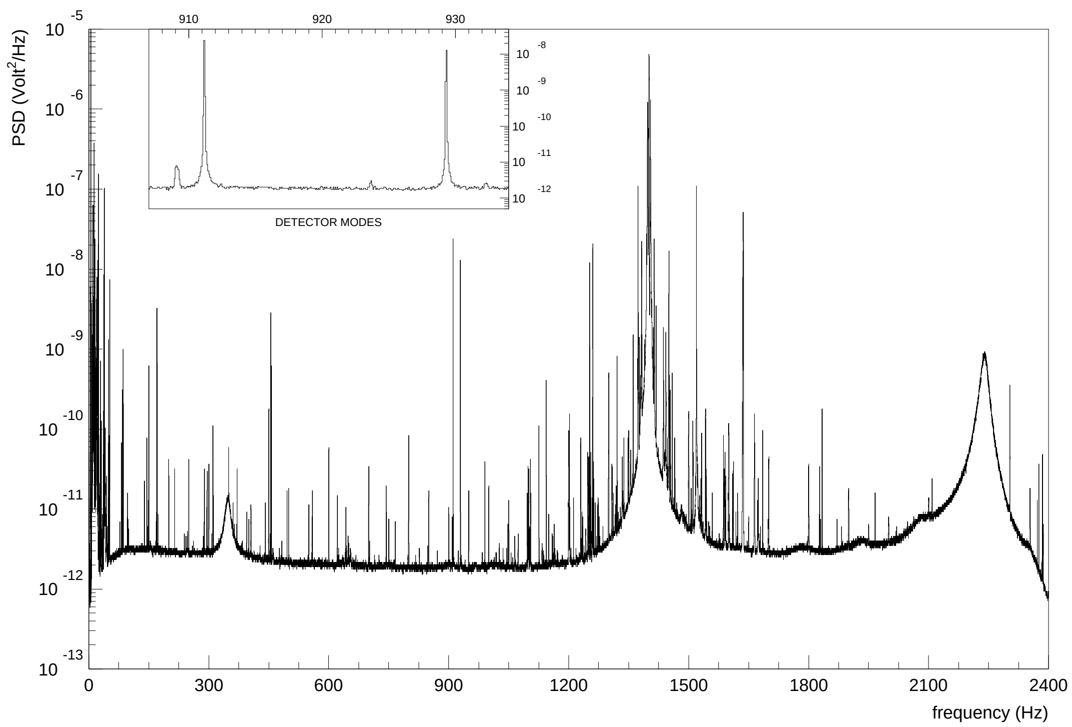
2.1 The real detector performances

The behavior of gw detectors now in operation can be described by the model presented in the previous sections, but only at a first degree of approximation. In fact, in some extreme cases, real detectors behave only as a pale shadow of the prototype. A proper data conditioning allows removing the unmodeled details in a consistent way.

2.1.1 From raw to WK filtered data

The model described in 1.3.1 is more or less likely to be verified in a real detector. In fact, the only strong source of wide band noise is certainly a Nyquist white noise source due to electronic noise. The other noise sources are narrow band, were "narrow" means 1 Hz at most, typically much less than this. Therefore, just by a matter of chance, it is unlikely that they directly disturb the narrow frequency band of sensitivity of the detector.

Figure 9 – A sample PSD of AURIGA data (UTC 1h 1 Aug 1999), with an enhanced view of the two normal modes of the coupled system



Of course one may ask what if actually an environmental narrow band noise is tuned to one of the sensitive modes: well, that's the time to take a screw driver and fix the detector hardware! Moreover, in the 1kHz region of the spectrum there used to be a lot of unstable resonance peaks, whose frequency could span in a daily cycle in a range of more than 10 Hz (see 4.2.3 for a closer view of some diagnostic data about AURIGA spectrum). Other peaks could appear and disappear during or after cryogenic (1K pot He refill) maintenance operations. One of the tasks of part of the group was in fact just to look day by day at the spectrum, trying to clean it up somewhat or at least to keep a good point.

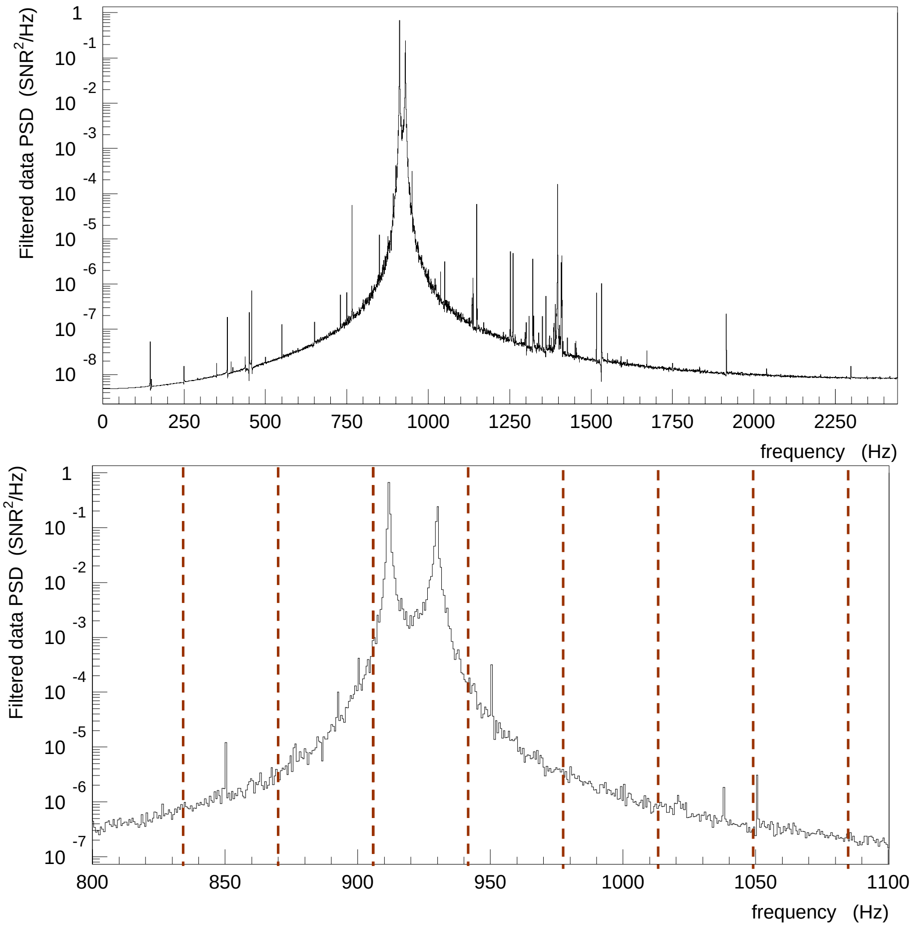


Figure 10 – (*top*) Spectral density estimate from a filtered data buffer (about 12s), whose integral is normailized to unity; (*bottom*) enhancment in the region near the sensitive modes. The dashed lines represent the sidebands that fold-up by aliasing with a decimation factor of 70 (35Hz final bandwidth).

During typical operating time, most of the power of unmodeled noise in the AURIGA PSD is concentrated far away from the sensitive bandwidth Figure 9. This let us apply the theory of WK filtering: as the signal-to-noise ratio function per unit Hz is highly suppressing everywhere else, the bias in the amplitude of the samples is negligible, and event search is minimally affected.

There is a third interesting stream of data, other than the raw detector output and the signal-enhanced WK filter output: the whitened data. Here we should remark that "whitening" we are speaking of refers only to the two fundamental resonant modes. The rest of the "whitened" spectrum is exactly as the raw data PSD. As we are not interested to the entire 2.4kHz band other than for diagnostic purposes, the data are decimated a factor 70 after the WK filter and therefore demodulated to a more convenient 70Hz stream *before* passing them to the whitening filter[8]. This is not a problem for aliasing, as the filtered data are narrowband. The whitened decimated data to a good approximation contain really only white noise.

Figure 11*c* shows a PSD of the data after the whitening filter applied to the decimated data. The fact that there are no remnants of the two modes of the system means that the model fully describes the noise of the system in the decimated bandwidth, which is a strict requirement in the recipe for a good WK filter.

[8] Obviously we are speaking of the whitening filter for the filtered data, which is not L but $L \circ W^{-1}$

2.1.2 Event search and spurious rejection

The data are now ready for event search. The **X** test statistic is computed using (1.29), and this quantity together with the amplitude, the SNR and ToA are the only information pertaining to the events recorded into the AURIGA database for offline consultation.

The database stores many other periodic data streams not directly related to the events, like the first four moments of filtered data distribution, or the residual correlation of the whitened data. They are used for diagnostic purposes, to record the status of the detector and for data quality validation: as we shall see in 2.2, they form the basic information of the veto system for AURIGA event list publication.

Figure 12 groups the main characters of the event list production for data exchange purposes: the events form a time series, which *only at this stage* is made through an explicit

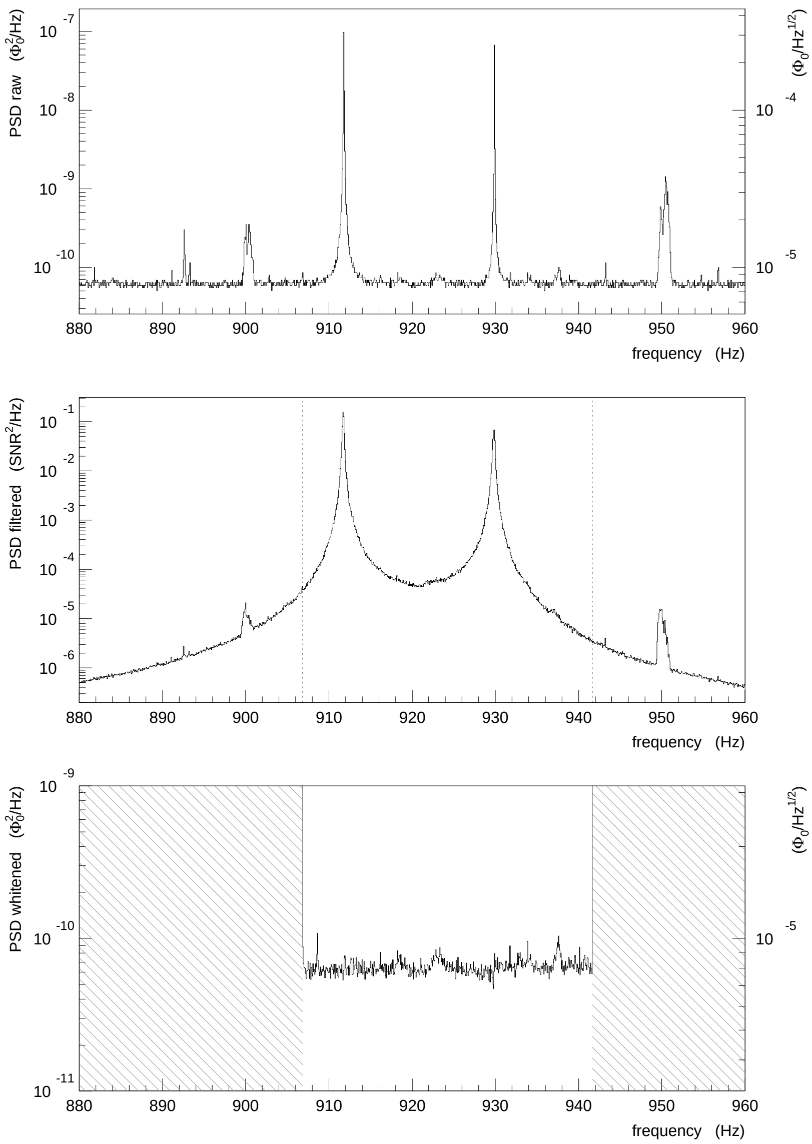
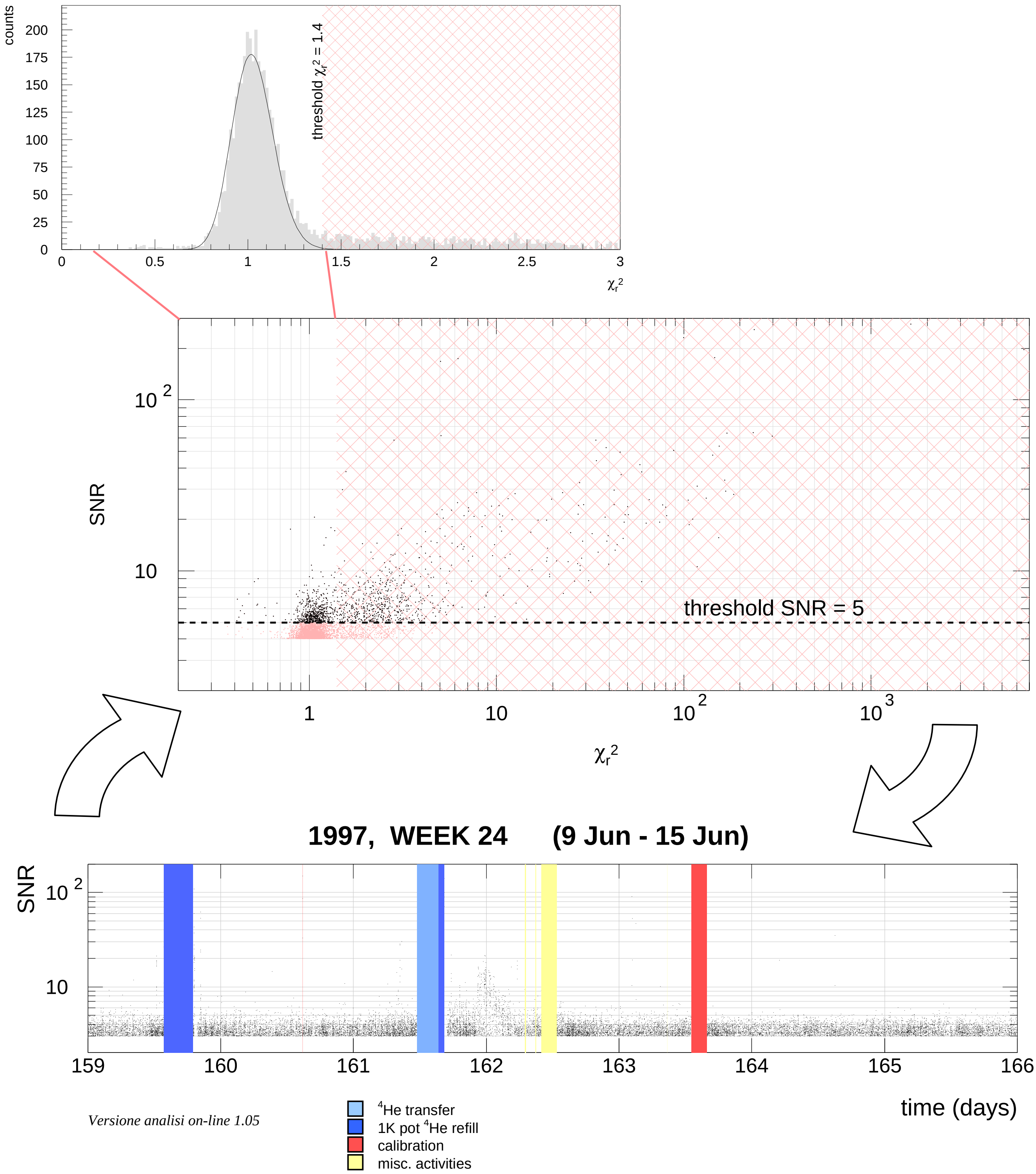


Figure 11 – In the power spectrum density of the raw data (*a*), enhanced around the detector's modes, the only small departures from the model of Eq. (1.43) are a few monochromatic disturbances, due to AC power sources. After the WK filter (*b*) the data are approximately band-limited, hence sub-sampling is allowed without aliasing. The power spectrum density estimate of sub-sampled whitened data (*c*) shows that the parameters of the noise model were estimated correctly.

threshold in amplitude in order to reduce the number of events to a rate which is useful for coincidence search (i.e. which gives raise to a tolerable false alarm rate). In fact, up to this point we let the max-hold event search saturate at its natural event rate, which is of the order of the reciprocal of t_w (or many *thousands* events per day). Keeping only the loudest events, with an adaptive threshold above SNR=5 the event rate decrease to about one *hundred* events per day. The event amplitude histogram (see Figure 8) is very steep, and a factor 2 more in event threshold would end up in a selection of just a couple of events per day. It is at this level that the $\mathbf{X}$ test statistic starts playing a role. In fact, almost all the events with SNR>10 are rejected, even with a relatively high false dismissal probability, about 10^{-4} (but this is what we expect from the theoretical χ^2 statistic, the actual efficiency is somewhat lower because the empirical $\mathbf{X}$ test statistic density function broadens somewhat due to SNR estimate uncertainty, filter biases and so on). This means that event characterization by goodness of-the-fit test is mainly targeted to a sound detection of high confidence level and good signal timing events, while it is ineffective at lower SNR, were the Gaussian bulk dominates.

Figure 12 – The event validation factory at work. (*bottom*) Event produced by the online analysis are vetoed by their time tag agaist interruptions of the normal operation of the detector because of maintence activity (see 2.3). (*middle and top*) To be released as candidate gw, an event is also thresholded in amplitude and its waveform must pass a single tailed χ^2 goodness of-the-fit test (acceptance set to an of 1-10-4).



2.2 Filter parameters estimate

Lock-in's are useful to estimate the central frequency and bandwidth of narrow spectral lines. An approximate feedback adaptiveness of the WK filter bandwidth is based on the whitened data power spectrum density.

Fast transients are the main problem with parameter estimation, but they are almost always associated with non-gaussianity of the data. On the other hand, when filtered data follow a Gaussian statistics it is likely that any other statistic test (like residual correlation in the whitened data) is just showing a mismatch of the estimated parameters with the true ones.

2.2.1 Overview of adaptive techniques

We have seen in 1.3.1 that only a bunch of parameters are needed for a parametric description of the detector noise and transfer function. These parameters have to be estimated to build the WK filter, but due to the unavoidable non-stationary behavior of the system some of their values changes significantly in time. If the non-stationarity is slow, i.e. occurs on a time scale longer than the relaxation time of the modes, the data analysis can track them by an adaptive algorithm. The parameters for the WK filter are updated with the ones from the slow tracking at the beginning of each hour.

For each parameter, an *ad hoc* procedure was applied [23].

- $\omega_k = \mathrm{Im}(p_k) \approx \mathrm{Im}(q_k)$ — the resonant mode frequencies are monitored by two fully digital lock-ins, implemented into the online data analysis software. The WK filter frequencies are practically identical to the pre-filtering one, so no new independent estimate is actually performed.
- $\Delta\omega_k = \mathrm{Re}(p_k)$ — the pre-detection bandwidth is known not to be a critical parameter, and is left to the value measured at the beginning of a data taking period[9]. It depends anyway on structural characteristics of the experimental setup, which cannot change without acknowledge of the experimentalist.
- $\Delta\omega_k^{opt} = \mathrm{Re}(q_k)$ — the WK filter (or *effective*) bandwidth is adapted so to keep flat the whitened data spectrum (see below)
- S_0 — the level of the amplifier's white noise is monitored by a lock-in displaced from the detector modes.[10]

The parameters are supposed to be slow-drifting, consequently we use moving averages to smooth their estimates on time scales of the order of the update schedule, which in turn is of the order of the system's proper relaxation time $(\Delta\omega_k)^{-1} \sim 10^3 s$, much longer than the WK

[9] It is also estimated online by fitting in the time domain the exponential decay pattern of the lock-in autocorrelation function. But this procedure has not been thoroughly tested yet.
[10] A recently implemented estimator seems to have greatly reduced variance and is endowed with a natural goodness-of-the-fit test to check for tilts in the spectrum or remove outliers from the estimate. You can see in 4.2.2 a preview of its features, and as a preliminary result it seems that S_0 is indeed a stable parameter during good operation periods.

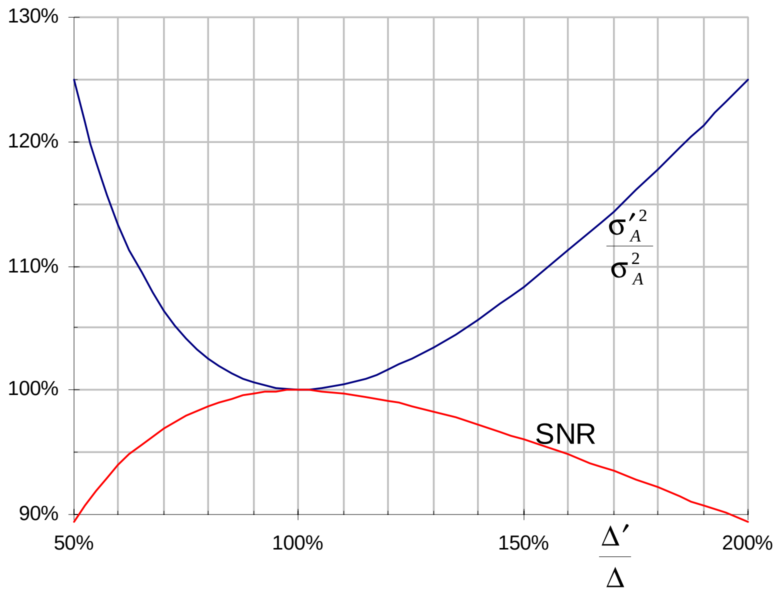


Figure 13 – Dependence of the variance of filtered data signal-to-noise ratio from the ratio of estimated (Δ') over true WK filter bandwidth (Δ). These results are correct in the hypothesis that the variation of S_0 are negligible. This point is delicate: in fact the bandwidth is determined (as pointed out in Figure 3c) both by the width of the detector resonace and by the level of the wide band noise. However, this parameter was monitored for a long time and shows to be constant irrespective to any other variation of the spectrum, except for hardware changes in the parameters of the electonics that control the amplifier, or because of unrecoverable failures in the electronic itself.

filter characteristic time $\left(\Delta\omega_k^{opt}\right)^{-1} \sim 10^3 s$. A non-stationary behavior faster than $\left(\Delta\omega_k\right)^{-1}$ does not allow estimating correctly the noise parameters, and therefore the analysis put a flag on consecutive buffers where this is frequently happens. Finally, a candidate signal would instead show up as a very fast variation of the detector output noise, limited in time and with a duration of a few times to the Wiener time.

Let us spend a few more words on the effective bandwidth of the detector, as it is the most sensitive and non-stationary parameter of the WK filter [36]. An error in its estimate shows up promptly in the whitened noise PSD. We exploit this effect in a feedback algorithm that corrects the parameters in order to keep the whitened PSD almost flat [23]. This algorithm usually converges in a couple of hours to the correct values (see 4.1.7).

Even if the algorithm has proved to be efficient in Monte Carlo simulations, it is heavily dependent on the assumption that there is only a mismatch in the parameters of the filter, and not a more serious problem with the noise model itself. Moreover, it may be useful to remark that when the whitening filter acts on a signal it does not whiten it as it does for pure noise –and it has to be so, or we would not see any signal at all! For this reason, a fast transient due to a large excitation of the system (even a gw event!) could produce in the hourly average a distorted whitened PSD, making us to think (wrongly) that an update of the effective bandwidth is necessary even if it is correctly estimated. So it is of primary importance to check the absence of signal in the data prior to the parameter estimate to avoid biases. The data buffers containing transients are selected and removed from the data stream used by the parameter-tracking algorithm.

The procedure is in principle simple, and is based on the observation that a high enough amplitude transient capable of sensibly wasting the hourly noise PSD spoils also the otherwise Gaussian noise distribution.

To detect this effect, the data are segmented in buffers about one minute long, and for each one the algorithm computes the kurtosis (i.e. the fourth moment of the distribution) for both filtered and whitened data, comparing it to the values one expects from a Gaussian distribution. The decision threshold we have chosen empirically corresponds to a confidence level at least 99% with respect to the fluctuation of the estimates obtained in a Monte Carlo simulation with purely Gaussian noise.

Effective temperature is the variance σ_{FLT} of the event amplitude estimate, i.e. of the filtered data (in the linear regime, see 4.3.2). It is related to the variance σ_{WHT} of the whitened data, and –if the system is following our parametric model– both are functionally dependent on the noise power spectrum density parameters: $\sigma_{WHT} = f(\sigma_{WHT}) = g(p_k, q_k)$.

Whichever the final choice of the best estimates (our attitude upon inconsistency of the different estimates was always to prefer σ_{WHT}), in order to estimate the standard deviation of the noise each buffer is checked for outliers, which are defined as everything in excess to 3 standard deviations. The procedure iterates recursively, computing again the standard deviation on the data left after outlier removal, until eventually the algorithm converges, or there are no more data left. This method is known as ‘Chauvenet procedure’, and its convergence is guaranteed most of the time for data distributed with Gaussian statistics. This gives as a corollary an independent check that the data buffer comes indeed from a Gaussian distribution: in case of poor convergence the buffer is not used for parameter estimate.

We further verify that there is no residual ‘color’ in the whitened data (i.e. its cross-correlation is not too far from zero). This is in the spirit that we are taking care automatically of perturbative corrections to the parameters, while bigger ones could be a symptom of more serious malfunctioning, and are usually handled manually. The parameter on which the analysis actually performs this test is a sort of derivative of consecutive samples.

A data buffer is considered to be Gaussian if its kurtosis is compatible with that obtained in Monte Carlo simulations within 99.7CL and if the Chauvenet algorithm converges within a few steps eliminating at most a few percentage of the data samples. Moreover, we require the correlation of the whitened data buffer not to exceed a 99.7CL threshold with respect to Monte Carlo simulations.

Sample pictures of good and bad buffers are reported in Figure 15. Figure 16 shows the result of this procedure for four hours of AURIGA operation.

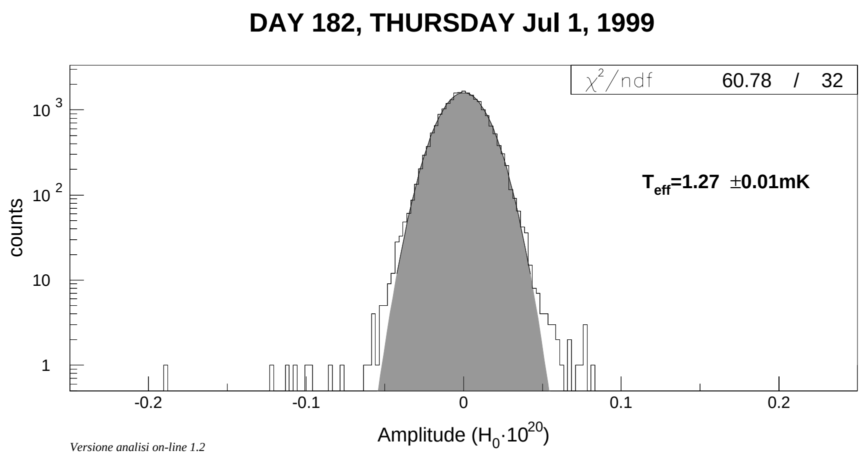


Figure 14 - Amplitude histogram of filtered data after during one day of satisfactory operation of the AURIGA detector. The data has been subsampled one per second to get almost independent samples. The small deviations from gaussian statistics at high amplitudes are due to the presence of signals ovrimposed on the noise. The variance of these data, expressed as effective temperature (T_{eff}), accounts only for the gaussian bulk. This graph is shown only as a teaching exampl. The valuse of T_{eff} actually used to estimate the SNR of each event is given for each decimated buffer (~2min), as a moving average of the last 10 buffers which passed gaussianity and whitening tests,

The discrimination between good and bad data buffers by requiring strict compliance of the noise with a parametric model is a cornerstone feature of the AURIGA data analysis. If one of the previous tests is not passed for a certain buffer, it is not considered in the computation of the parameter correction, and the parameter update is delayed. This doesn't affect the search for candidate gw events, that is performed anyway also in this buffer. However, a long sequence of 'bad' buffers is considered unreliable also for gw detection, and marked for removal from duty cycle.

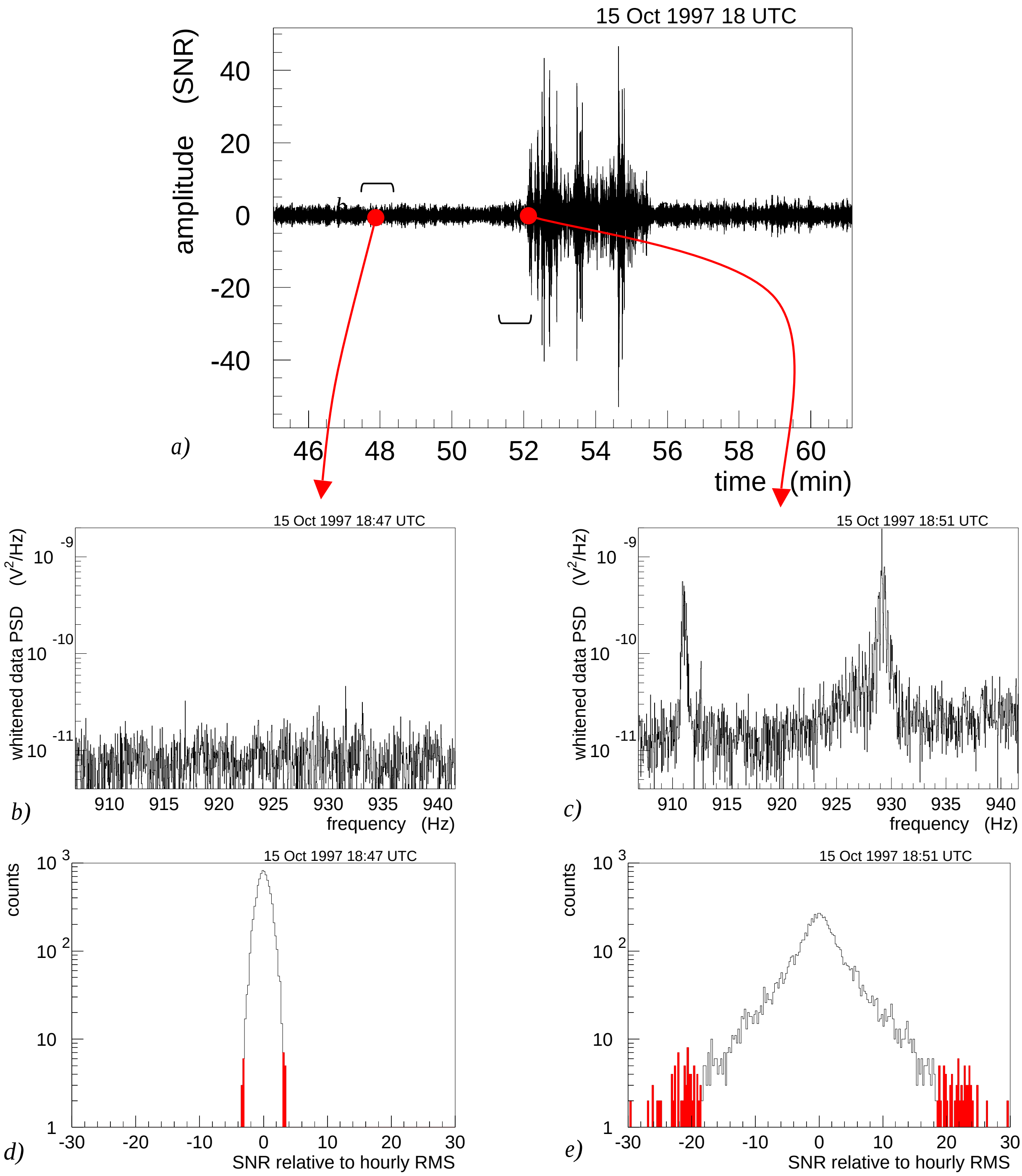


Figure 15 - The filtered data (*a*) are divided into buffers of 2 minutes. The two ones marked with brackets have quite different statistical distributions (resp. *b* and *e*), particularly on tails beyond 3 times the Root Mean Square (*solid color*). The non-gaussian buffer is not let enter at all the effective noise temperature estimate (which is a RMS moving average). Notice that, for the same 'bad' buffer, the whitening filter seems no more working properly (see *c* and *f*), in particular it mimics a displacement of the real part of the zeros q_1 and q_2 (see 4.1.7).

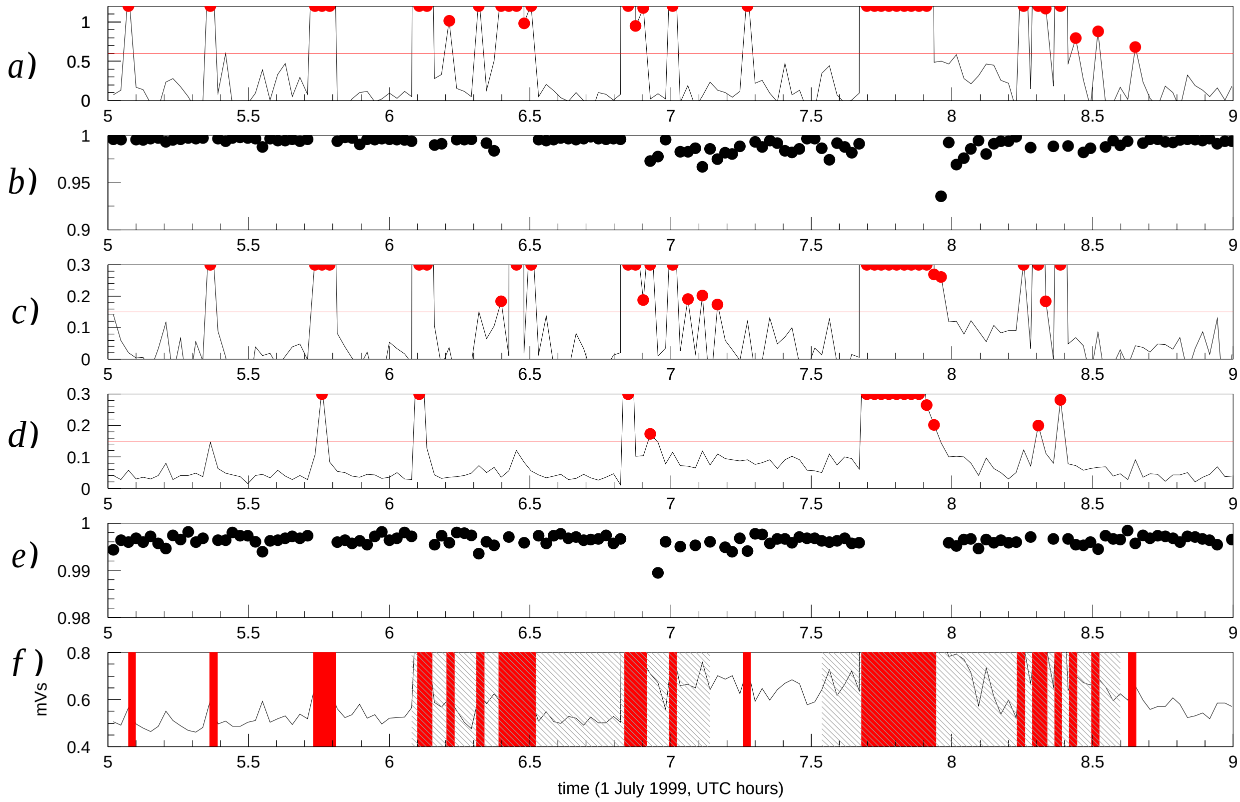


Figure 16 – An overview of the results of standard statistical tests performed on filtered and whitened data during four hours of AURIGA operation (UTC $5^h \div 9^h$ 1-Jul-1999): kurtosis statistical test for WK-filtered and whitened data (*a,c*); the residual fraction of WK-filtered and whitened data kept after the Chauvenet selection (*b,e*); the correlation of whitened data (*d*). Thin horizontal lines are the acceptance threshold, equal to three times the outcome of the analysis with pure gaussian nose simulated with an ARMA model. (*f*) The bottom graph shows: 1) the variance of the filtered data buffers (*thin line*) in units of effective temperature; 2) the time periods corresponding to data buffers with bad statistical properties (*solid red banners*); 3) the vetoed periods of operation due to too frequent bad data buffers (*dashed areas*).

Two main situations may then arise: either there is a dominant contribution of the modeled quasi-stationary noise with short time periods showing unmodeled excess noise and/or signals, or the data are dominated by unmodeled excess noise. In the first case, the bad buffers are rare enough so that the analysis is able to reliably estimate the noise parameters and to adapt the WK filter to any slow non-stationary noise behavior by using good buffers only. The estimate of the noise parameters is therefore performed in a reliable way by using only the periods when the modeled noise is dominating and disregarding the bad buffers, which instead contain some “signals”.

This is therefore a satisfactory condition for detector operation, corresponding in Figure 16*f* to the time periods not covered by the dashed pattern. Figure 14 shows a sample amplitude histogram of the filtered data during the satisfactory operation of the detector during the same day of Figure 16; the statistics is Gaussian with small excess counts in the tails due to the “signals” present in the rare bad buffers.

The other main operating condition, i.e. that the data are dominated by unmodeled excess noise, is identified –as a rule of thumb – when more than 4 buffers within a fixed time window 10 buffers long failed the tests. In this condition, the WK filter is badly matched to the noise –in fact it could be that WK linear filter theory is not applicable. The analyzed data therefore lack of self-consistency and the output data are vetoed, as shown in the bottom graph of Figure 16 with dashed areas. Under this condition the detector is not necessarily blind and with advanced different noise models and/or parameters one could recover some of the vetoed observation time.

2.3 Duty cycle estimate

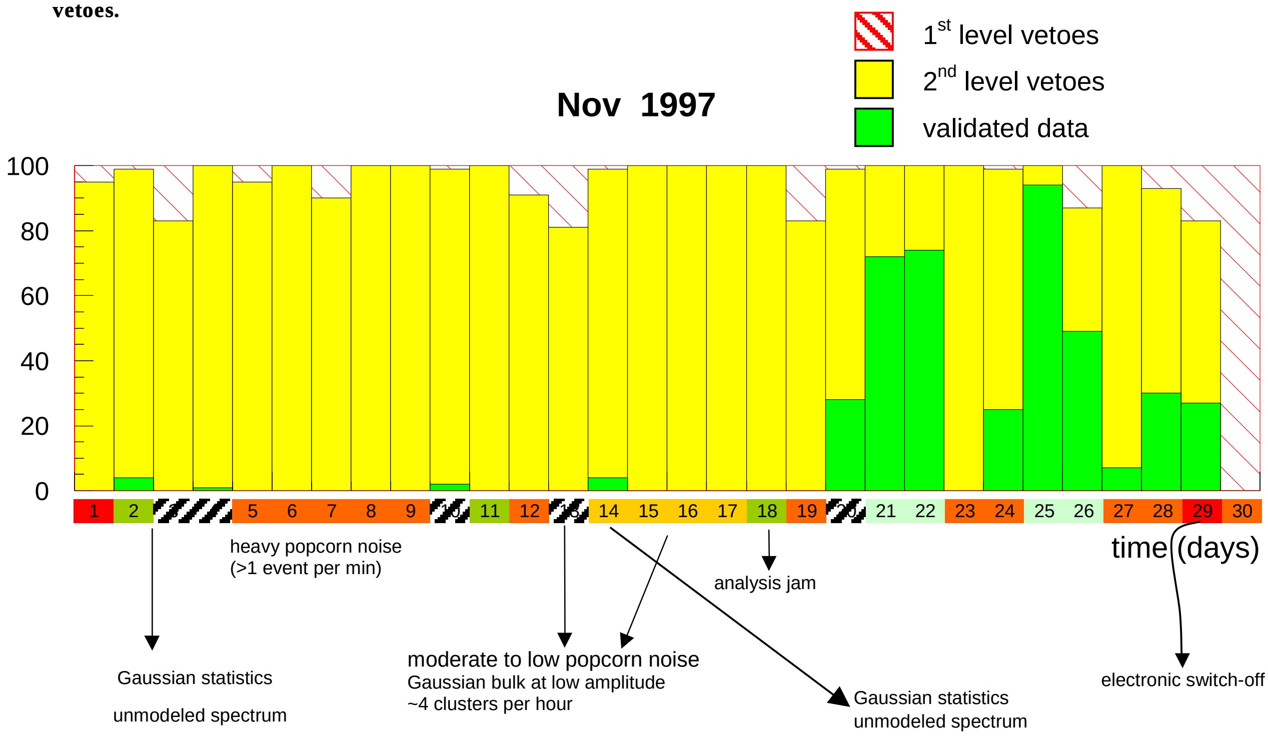
A period of 'good operation' is defined as one in which detection efficiency is close to 1, and the biases on amplitude and time of arrival estimate are negligible. It has to be long enough for the estimators of the filter parameters and of the sensitivity to converge. Eventually it should be judged on the capabilities of consistent event detection, at least for simulated events.

2.3.1 Statement of the problem

The operative time of the detector –often referred to as *duty cycle*– is shortened by maintenance operations or failures in any of the many aspects of the detector: cryogenics, vacuum, electronics. Even seismic activity in Afghanistan or in the central Italy in 1997 was registered as series of huge events, which blinded the detector for a few minutes. All these kind of disturbances have one common characteristic: they can be well delimited in time and their origin is known. So the usual procedure is to create a list of time intervals, with duration from 1 minute upwards, and consider the detector not operative inside them. We shall call them *first level vetoes*.

There is a second kind of more annoying disturbances, that can arise from sources that are not monitored even indirectly –e.g. from leakage in the needle valves in the internal liquid ^{4}He temperature stage of the dilution refrigerator. They show up in the data in a large range of behaviors: a series of apparently random short excitations, or non-stationary resonance modes near the sensitive frequency range of the detector, or a long lasting non-Gaussian

Figure 17 – The placement of 1st and 2nd level vetoes on a sample month, with some manual annotations, from the logbook of AURIGA data validation. In the line below the graph the *orange-red* stands for "irrimediable" situation (data that can be 1st level vetoed *a posteriori*), the *green* for data tagged for manual re-analysis, the *yellow* for data that could be recovered with appropriate advanced analysis tools. Notice the *"red"* on 29th: the entire day sould have been vetoed, and validated data are in fact a perfectly Gaussian... switched off amplifier! This is the most common mistake found in placement of manual vetoes.



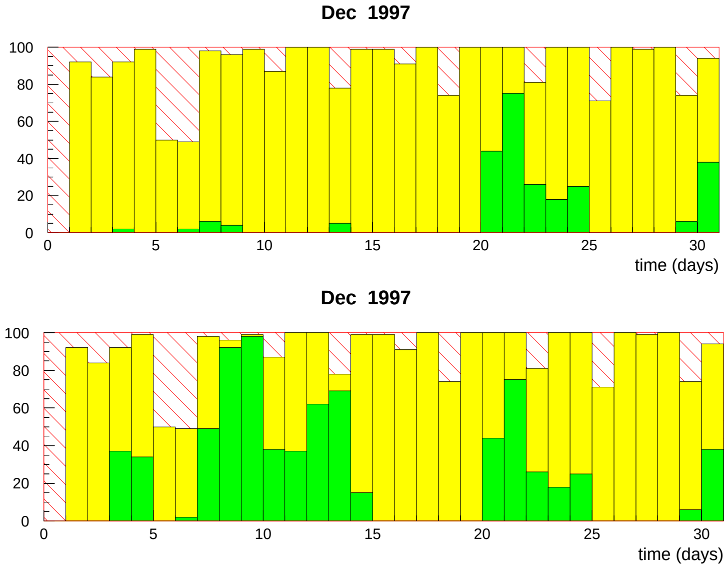


Figure 18 – Exercise: Dicember 1997 with 'wild' running analysis and after manual correction, and re-analysis of the data. The operative time of the detector goes from 8.3% to 25%.

excess noise.

A big effort has been spent to remove the sources of these phenomena, but the investigation has not reached final result, and will eventually be carried on with the help of a test facility now being set-up inside the building housing the AURIGA experiment. As regards the data already acquired, we have devised a validating procedure to highlight and eventually remove all periods in which the noise of the detector is undoubtedly affect by unknown disturbances, as we shall describe below. We shall call these removed time intervals *second level vetoes*. Let us summarize the chief characteristics of both two kinds of vetoes.

1st **Level vetoes** (from logbook of laboratory operations, also on the Web)

- Definitive (must be checked carefully)
- Independent from data analysis software release
- No model

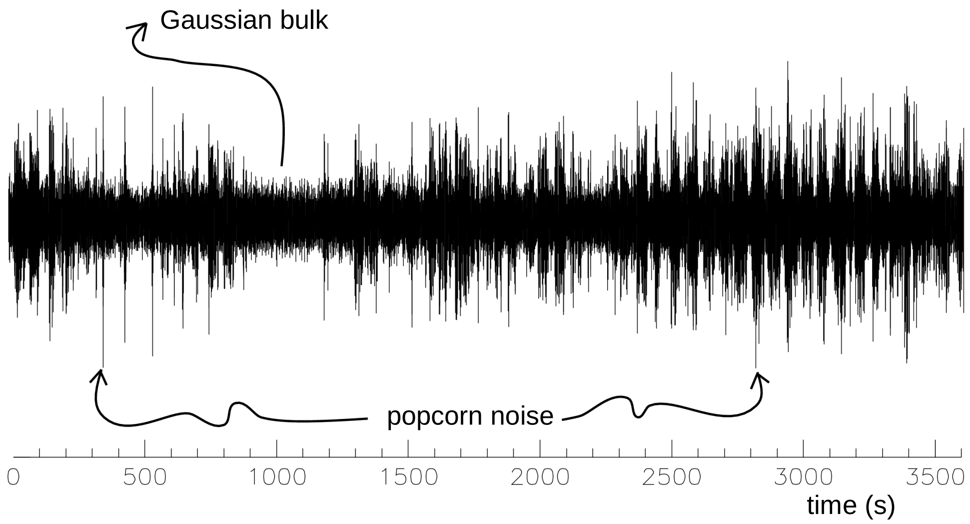
2nd **Level vetoes** (automatic: 'Chauvenet' criterion, Gaussian tests, residual correlation after whitening, …)

- Depend on the analysis and on the model hypothesis (es: '$<3\sigma$')
- Easy to implement (=automatic)
- Limited by the efficiency of the parameter-tracking algorithm

With a rigorous application of all vetoes the final duty cycle is very low (around 30%). Three are the main reason for this:

1. *analysis jam*: the analysis was running with non up-to date parameters, outside the range of the feedback algorithm. This was true in particular in the early days, when off-line analysis was let running unattended, without manual status check (see Figure 20). There is in fact a major objection with the adaptive algorithm described in 2.2: it is not stable for large oscillations of the parameters. When the deviation of the parameters from optimal configuration is small, the test statistics still agree in confirming that the period is good, hence the correction to the parameters is computed and applied for the next hour. But there is a critic threshold above which the validation test fails only because the mismatch in the filter parameters was

Figure 19 – Popcorn noise in a one hour buffer of filtered decimated data.



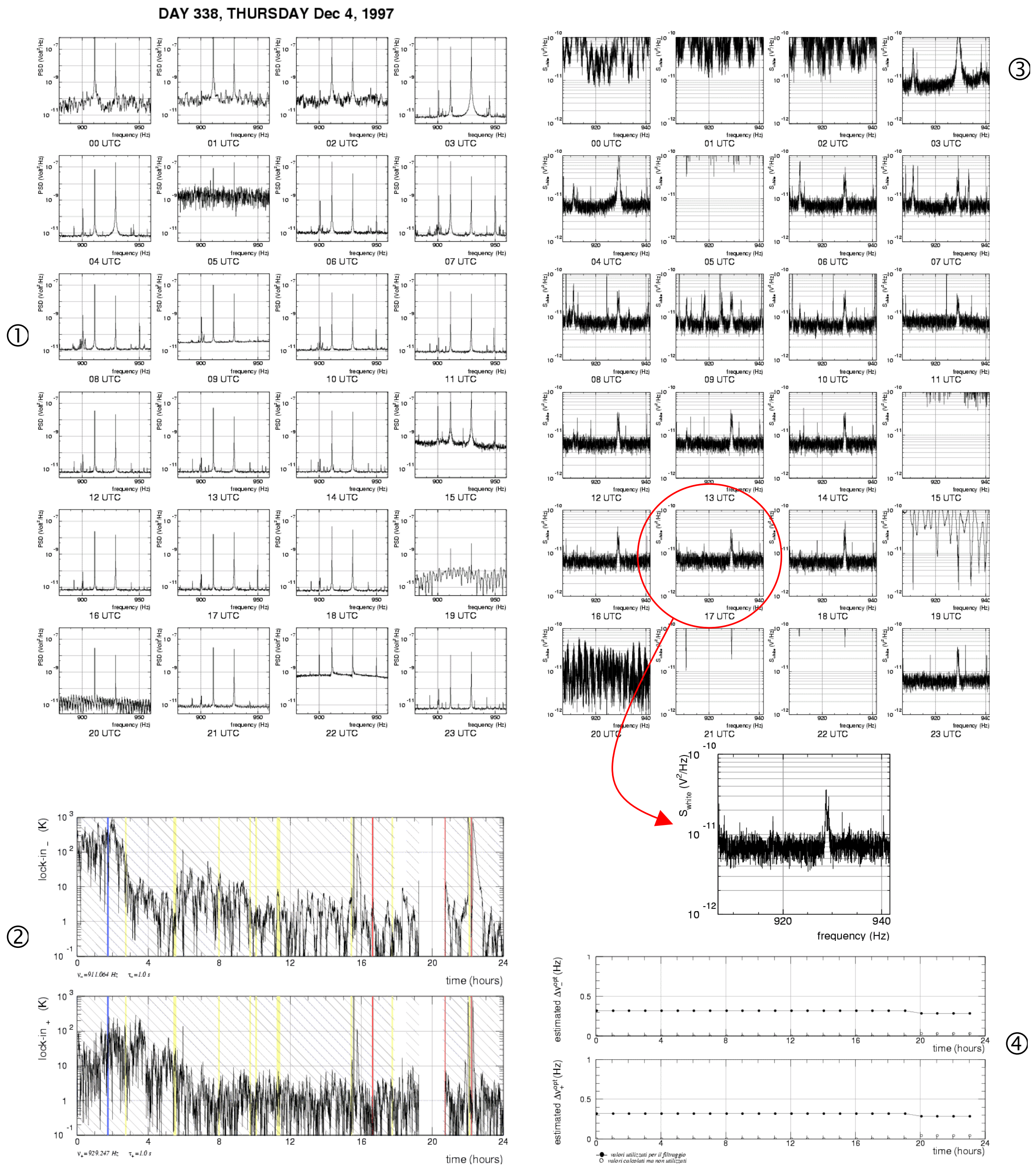
too high. The system enters a vicious loop: without update, the parameter drifts are not followed, and as the validation test keep failing, the system eventually recovers only if the parameters by chance enter again the allowed small deviation range. Exceptions apart (see Figure 18), the typical gain in analyzing all data with manual control is about 10÷15% with respect to wild running analysis.

There is another kind of "analysis jam" due to periods during which the amplifier is switched off, but the data are anyway acquired and analyzed. All Gaussian tests promote these periods to "good operation". If they last too long, eventually the parameters are updated to a no return region: when the system recovers, the "colored whitened spectrum" (see below) problem prevents any subsequent parameter update. This is a known issue, these periods fall in the category of 1st level vetoes, but the present stable version of the analysis does not relate parameter update control to the list of vetoes.

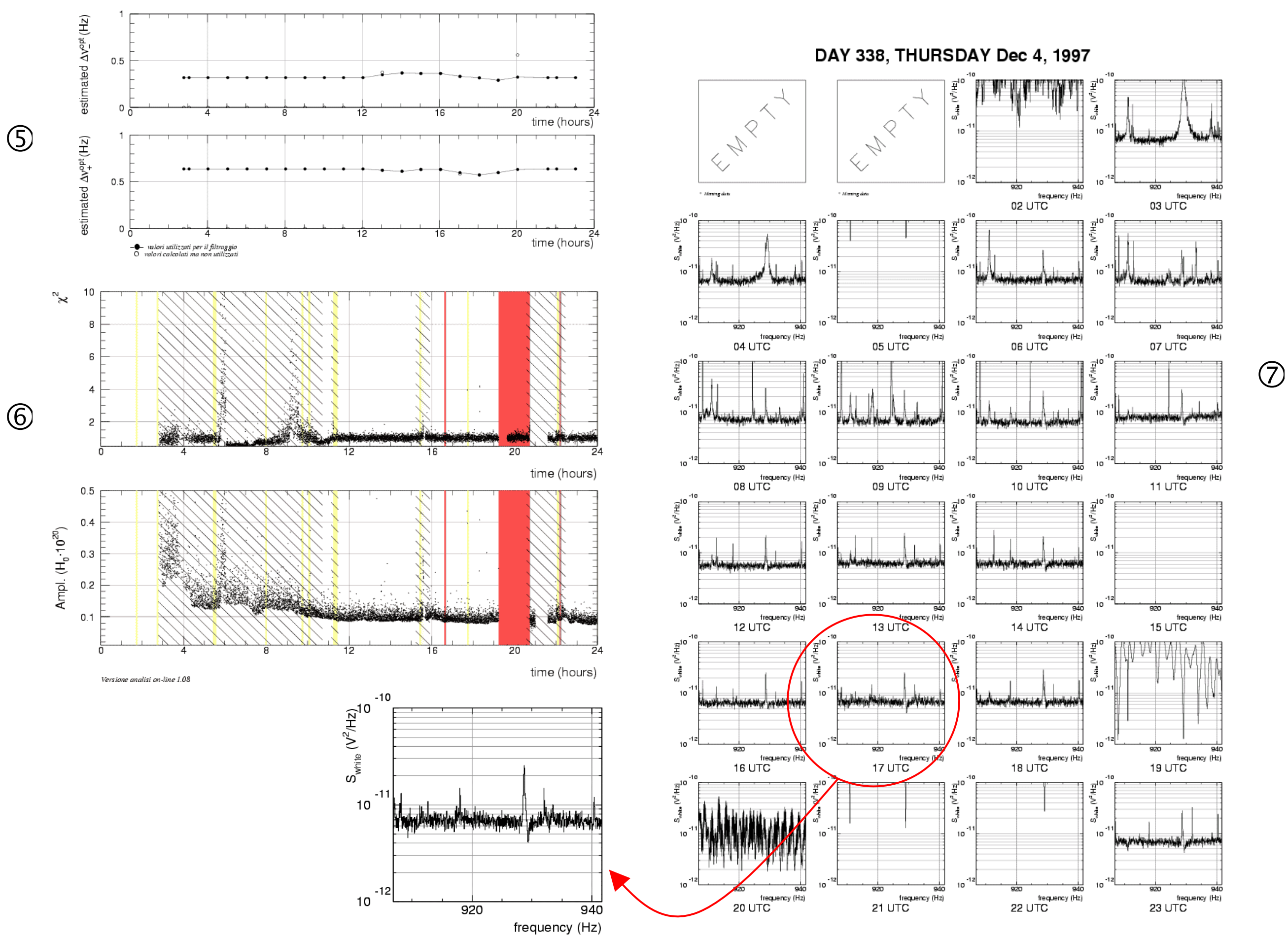
2. *colored whitened spectrum*: the spectral density of whitened data is corrupted by residual unmodeled peaks. It is not unlikely that the sensitive frequency interval near the modes is rather clear, which means that the signal amplitude, the standard deviation of the noise, and perhaps also the timing errors are not completely spoiled, because the unmodeled peaks are suppressed in the filtered data. The problem is that the **X** test statistic (which is computed over the full decimated whitened frequency interval) is no more a χ^2 test and there is a big chance of false dismissal. Also the variance of the whitened data is biased, therefore it cannot be used as an alternative estimate for the amplitude variance[11]. The choice of removing from the exchanged data those whose noise PSD is not strictly fitted by the ARMA model is responsible for 5÷10% of duty cycle shortage.
3. *popcorn noise*: The system is indeed overloaded with unmodeled noise, typically as clustered events, from 100 to 4 per hour. This happens for 20÷30% of duty cycle, and is often associated with cryogenic operations/failures.

[11] If the constraint on the residual color of the whitened spectrum is relaxed, then another issue opens, as for consistency also the parameter update should neglect this test.

Figure 20 – Chronicle of an unusual analysis jam discovered by manual check. ① The spectral density of the detector near the modes indicate that at the beginning of 4th Dec 1997 the system was disturbed, and than it recovered during the rest of the day. ② The mean energy of the two sensitive modes measured with numerical lock-ins confirm with finer detail this fact. The dashed gray shadows represent 2nd level vetoes proposed by the analysis, and they cover the entire day (the clean period about UTC 20h is a period with switched-off electronics, whose corresponding 1st level vetoes was not set). ③ Whitened data PSD at each hour the same day; the rather pronounced peak in the correspond to a permanent mismatch in the sensitive bandwidth of the "+" mode. ④ The bandwidth of the two modes, hour by hour, shows clearly that this parameter is freezed for the entire day



After this discovery, the automatic offline-analysis had to be stopped, tape rewinded, and the parameters manually set to the correct initial values. ⑤ The slow drifts of the estimated bandwidth in both modes after UTC 12^h is a hint that now the engine is working. ⑥ Temporal plot of the Χ test statistic and amplitude for each event above SNR=5 found in the new analysis run. About 8 hours appear now unvetoed, within which the background events follows the correct statistics (the red stripe shows that the electronic failure was now correctly set as a 1^{st} level veto) ⑦ A check to the whitened data PSD in the new offline analysis run reveals that there is still a peak in correspondence to the "+" mode (the reduced variance of the estimate is an artifact due to a different binning). What is happening? Well, checking carefully an enhancement of the spectral estimates of the previous run, it is clear that there was a second unmodeled peak next to the "+" mode one. Actually, in the new run the whitening filter overcompensate the "+" mode peak, digging a hole near the spurious peak!



It is possible that part of the vetoed time could be recovered tuning the minimum duration of an operative run. In fact, if the problem is to falsify the hypothesis of a gw burst in coincidence with a specific trigger, than even periods like that depicted in Figure 19 (for instance in the time interval around 1000s) are worth a look. But before we can systematically release such kind of data a few more investigation has to be performed on the final non-Gaussian statistical methods to deal with them.