Search for gravitational wave bursts by wavelet packet decomposition: The detection algorithm

M. Camarda

Dipartimento di Ingegneria dell'Informazione, Via Gradenigo, 6/B, 35131 Padova, Italy,
and INFN Laboratori Nazionali di Legnaro, I-35020 Legnaro, Padova, Italy

A. Ortolan[*]

INFN Laboratori Nazionali di Legnaro, I-35020 Legnaro, Padova, Italy
(Received 21 June 2006; published 19 September 2006)

We present a novel method based on wavelet packet transformation for detection of gravitational wave (gw) bursts embedded in additive Gaussian noise. The method exploits a wavelet packet decomposition of observed data and performs detection of bursts at multiple time-frequency resolutions by the extreme value statistics. We discuss the performances of detection algorithms (efficiency and robustness) in the general framework of hypothesis testing. In particular, we compare the performances of wavelet packet (WP), matched filter (MF), and power filter (PF) algorithms by means of a complete Monte Carlo simulation of the output of a gw detector, with the detection efficiencies of MF and PF playing the role of upper and lower bounds, respectively. Moreover, the performances of impulsive filter (IF) algorithm, widely used in the data analysis of resonant gw detectors, have been investigated. Results we get by injecting chirplet signals confirm the expected performances in terms of efficiency and robustness. To illustrate the application of the new method to real data, we analyzed a few data sets of the resonant gw detector AURIGA.

I. INTRODUCTION

Detectable gravitational wave (gw) bursts are expected to be produced by compact astrophysical sources, such as core bounce in SN collapses, merging of neutron stars and/ or black holes [1], and by other high-energy cosmic phenomena, for instance the inner engines that power gamma-ray bursts [2]. Up to now, due to the absence of gw detection, the characteristics of burst signals are largely unknown, apart from some theoretical and numerical calculations of black hole ring-downs [3] or cusps and kinks in a primordial network of cosmic strings [4]. In general, the understanding of impulsive gw sources is poor; in particular (i) the calculations and numerical simulations of core collapse agree on the gw weakness, but the expected energetics spans many orders of magnitude (from 10^{-8} to $10^{-2}M_\odot c^2$) [5]; (ii) the gw burst rate is highly uncertain e.g., from optical observation of SN explosions in the nearby galaxies [6], we can estimate 1 SN event every 30–50 yr per galaxy; however, the electromagnetic radiation may not trace the gw luminosity; (iii) gw waveforms are not yet accurately modeled as core collapse and/ or black hole formation are complex physical phenomena, that require the theory of general relativity in its most fundamental (i.e. nonlinear and singular) structure.

Moreover, on the experimental side, we must face technological limits of operating gw detectors, either resonant or interferometric, that constrain their actual sensitivity to $10^{-21} \le h_{\min} \le 10^{-19}$, where $h_{\min}$ is the minimum gw burst amplitude detectable at unitary signal-to-noise ratio

(SNR) [7]. The wide excluded region in the rate vs gw burst amplitude plane, already set by the IGEC collaboration [8] and the LIGO observatory [9], confirms the challenging task of the present generation of detectors. Hence, in the hope of detecting gw bursts, we must fully exploit the detector sensitivity by means of suitable digital signal processing techniques. From this point of view, the transient gw signals have peculiar characteristics, namely: (i) sources and waveforms are unverifiable as gws are irreproducible in an Earth laboratory; (ii) astrophysical events capable of producing measurable gw could be very rare; (iii) the arrival time of gw signals is unknown; vi) SNRs could be (very) low. Presumably, gw transient signals have a short duration (≤ 1 s), but there is no specific model (or template) to fit the experimental data.

By *detection* of a transient signal we mean the identification of its presence in additive Gaussian noise. A detection algorithm, also known as "event trigger generator" (ETG), consists in the selection of candidate events or "triggers" in the continuous data stream of a detector output. The general approach for the gw detection is based on the thresholding of a suitable statistics. The statistics of the detection algorithm can be computed from the output of a filter matched to the template of the incoming signal (MF) [10] or to the impulsive response of a resonant detector [impulsive filter (IF)] [11]. For the detection of signals of known duration and frequency band, the output of power filters, designed to localize the signal power in the time-frequency plane, can be used as a detection statistics [12]. The power filter (PF) algorithm is based on the power filter for all frequencies within the detector bandwidth. The application of time-frequency decompositions of detector

*Corresponding author: ortolan@lnl.infn.it

data to gw searches is relatively recent [13]. In particular, the WaveBurst method [14,15] is based on a nonparametric statistic obtained through a bank of wavelet packet filters and a clustering of higher coefficients. For an exploration of a variety of other statistics, see Ref. [16]. Here we present a detection algorithm for gw bursts with a different statistics, which is based on the maximum coefficient of wavelet packet decomposition of the observed data. The statistics make no *a priori* assumptions about waveforms, time of arrival, duration, or spectral contents. The algorithm involves the separation of the detection problem from the estimate of the signal characteristics, e.g. amplitude, time of arrival, duration, spectral contents, etc. For instance, in wavelet analysis, the separation of detection—from estimate—phases can be achieved by introducing two independent thresholds: one for the detection algorithms, in order to control the false alarm rate, and the other for denoising (or better estimating) the signal waveform [17]. With the separation of the detection and estimate phases we have, on one hand, that data that are classified as signals (because the null hypothesis is false) are further processed in order to estimate the waveform characteristics in the presence of noise. On the other hand, with a given level of probability, the remaining data are correctly interpreted as pure detector noise. In this paper we focus on the detection algorithm whereas the signal estimate is addressed to a forthcoming paper [18]. The theoretical framework of the detection problem is the classical Neyman-Pearson theory of hypothesis testing [19].

The plan of the paper is as follows. In Sec. II we give a brief introduction to the wavelet theory and its connections with multiresolution analysis. Section III is devoted to the illustration of some detection algorithms. In Sec. IV we describe the wavelet packet algorithm and its potentialities for the gw bursts detection. In Sec. V the performances of four detection algorithms, based on matched filter (MF), impulsive filter (IF), wavelet packet transform (WP), and power filter (PF), have been evaluated by means of a complete Monte Carlo simulation of signals and noise of a resonant gw detector. A few data sets of the AURIGA detector have been analyzed to show how the method works with real data. In Sec. VI a conclusion is drawn and the future research potentials of WP algorithm (in particular its extension to a network of gw detectors) are given.

II. WAVELET AND WAVELET PACKET ANALYSIS

In digital data processing, mathematical transformations (either linear or nonlinear) are applied to the sampled output of a detector to obtain further information that is not readily available in the raw samples. The wavelet transform is a suitable mathematical tool that allows enhancement and recovery of time-frequency structures that would otherwise remain hidden in the noise.

The *continuous wavelet transform* decomposes the signal on basis functions, obtained by translating and scaling a unique prototype $\psi(t)$, called mother wavelet. The basis functions are waveforms of limited duration (the opposite of the sinusoidal basis of the Fourier analysis) which allow both time and frequency localization of the signal spectral components. This leads to their key feature: wavelets give efficient representations of signals which exhibit transient behavior [20]. The basis functions

$$\psi_{s,\tau}(t) = \frac{1}{\sqrt{|s|}} \psi\left(\frac{t-\tau}{s}\right) \qquad s, \tau \in \mathbb{R}, s \neq 0 \quad (1)$$

can be obtained by varying the translation and scale parameters τ and s which, in turn, change the temporal and spectral extension of the wavelet. Therefore, scaling of $\psi(t)$ allows the partition of the time-frequency plane at different resolutions. Moreover, if the mother wavelet $\psi(t)$ satisfies the admissibility condition $\int_0^\infty |\Psi(\omega)|^2/\omega \, d\omega < \infty$, we are able to reconstruct exactly the original signal $x(t)$ from its continuous wavelet transform [21].

The *discrete wavelet transform* (DWT) is defined, without loss of information, by the discretization of s and τ on a dyadic grid, $s = 2^{-j}$, $\tau = n2^{-j}$ with $i, j \in \mathbb{Z}$. Mallat and Meyer [22] discovered a fundamental relation between the discrete wavelet analysis and the multiresolution analysis (MRA) i.e. the construction of function approximations in various closed subspaces $V_j \subset L^2(\mathbb{R})$ associated to the scale $s = 2^{-j}$. In the framework of MRA, one also defines the *scaling function* $\phi(t)$, orthogonal to its discrete translations $\phi(t-n)$, which produces an orthogonal basis in each subspace V_j:

$$\{\phi_{j,n}(t)\}_{j,n\in\mathbb{Z}} = \{2^{j/2}\phi(2^j t - n)\}_{j,n\in\mathbb{Z}}. \quad (2)$$

The approximation a_j of a signal x at the level j is defined by the orthogonal projection of x on V_j with respect to the basis $\{\phi_{j,n}(t)\}_{j,n\in\mathbb{Z}}$.

The scaling function $\phi(t)$ can be built on a low-pass FIR filter $h_0[n]$ called conjugate mirror filter (CMF) [20]. By construction, the CMF h_0 and the corresponding high-pass FIR filter h_1 [constructed by reversing and multiplying by $(-1)^k$ the coefficients of h_0: $h_1[n] = (-1)^n h_0[-(n+1)]$] are half-band filters that satisfy the power complementarity condition $|H_0(\omega)|^2 + |H_1(\omega)|^2 = 1$. The mother wavelet $\psi(t)$ can be obtained from the scaling function $\phi(t)$ and the high-pass filter h_1. One can demonstrate, with minor conditions on $h_0[n]$ [23], that the wavelet $\psi(t)$, scaled and translated, gives rise to an orthogonal basis of the orthogonal complement of V_j in V_{j+1} that we call W_j:

$$\{\psi_{j,n}(t)\}_{j,n\in\mathbb{Z}} = \{2^{j/2}\psi(2^j t - n)\}_{j,n\in\mathbb{Z}}. \quad (3)$$

Let $\{a_{j_0,n}\}_{n\in\mathbb{Z}} = \{\langle x, \phi_{j_0,n}\rangle\}_{n\in\mathbb{Z}}$ (i.e. the orthogonal projection of x over the basis functions $\phi_{j_0,n}$) and $\{d_{j,n}\}_{n\in\mathbb{Z}} = \{\langle x, \psi_{j,n}\rangle\}_{n\in\mathbb{Z}}$ ($j_0 \leq j < 0$) (i.e. the orthogonal projections of x over the basis functions $\psi_{j,n}$) the *approximation* and

detail coefficients, respectively. We can decompose any signal $x(t) \in V_0$ in a unique linear combination of scaling functions and wavelets by means of its approximation coefficients $\{a_{j_0,n}\}_{n \in \mathbb{Z}}$ at the level j_0 and detail coefficients $\{d_{j_0,n}\}_{n \in \mathbb{Z}}, \ldots, \{d_{-1,n}\}_{n \in \mathbb{Z}}$ from level j_0 to -1. The complete set of approximation and detail coefficients represents the DWT.

The multiresolution theory plays a crucial role in the implementation of the DWT as it provides a fast algorithm (analogous to fast Fourier transform) that avoids unnecessary calculations of mother wavelets and scaling functions. Such a decomposition algorithm—known as the *Mallat algorithm*—is based on a classical two channel subband coder implemented by the CMF bank. The DWT is obtained by iterating the CMF and the decimation by a factor 2 on approximation coefficients of the previous level. The resulting computational complexity depends linearly on the number of coefficients of the starting level.

However, wavelet analysis is not appropriate for all signals. In fact, the characteristic logarithmic frequency resolution is no more efficient when signals have high frequency spectral contents (e.g. bandpass signals). *wavelet packet* [20] analysis provides instead a "fine" frequency resolution also in high frequency bands. The idea of Coifmann, Meyer, and Wickerhauser [24] was to extend the iteration of CMF bank not only the approximation coefficients but also on the detail coefficients.

Wavelet packet transforms can be univocally represented by an *admissible* binary tree, where open nodes (the leaves) identify the associated partition of the time-frequency plane. Nodes of an admissible binary tree have either 0 or 2 branches.

The vectorial subspaces represented by open nodes of an admissible binary tree are mutually orthogonal and their union is equal to the original space, identified by the root of the tree. It can be shown [20] that any admissible binary tree identifies an orthogonal wavelet packet basis in the space V_0 and a different partition of the time-frequency plane.

III. DATA ANALYSIS

A. The search for gw bursts

The sparsity of signals (and their relative weakness) entail that the output of a gw detector is normally dominated by its intrinsic noise sources, as prescribed by the fluctuation-dissipation theorem [25]. The superposition of main intrinsic noise sources, e.g. thermal, backaction, and wideband noise due to Brownian motion and amplifiers, respectively, forms the detector sensitivity curve and the bandwidth available for gw searches. In addition to the intrinsic noises there are spurious disturbances, originated by many environmental noises (e.g. seismic noise, electromagnetic interferences, cosmic rays, etc.), with amplitudes incompatible with spontaneous fluctuations of the system.

Clearly, such "excess noise sources" (or more correctly spurious signals) may jeopardize the true gw signals. Unfortunately, predictions on gw burst rates and amplitudes are widely spread [26]: they are typically large enough to have a reasonable hope of observing something during the next years, but small enough that the probability of missing a signal when it occurs must be reduced as much as possible. In order to attain a robust and confident detection of gw bursts with an arbitrarily low false alarm rate (e.g. 1 event per hundred years as required by Supernova Early Warning System [27]), we have to take into consideration four steps of paramount importance:

(a) *Data conditioning algorithms* that identify the time intervals and the frequency bands available for the gw search. The aim of these algorithms is a careful removal of the instrumental effects either by discarding the maintenance periods, hardware malfunctioning, seismic or acoustic excitations, etc., or by identifying the frequency bands contaminated by environmental disturbances. Data conditioning may also include a whitening filter designed to remove the correlations among the output samples.

(b) *Detection algorithms* that assess the probability of a data set to contain spontaneous fluctuations or signals (gw bursts and/or local spurious disturbances); of course, the detection is performed during the on-times and within the detection band. At this stage of analysis, signals due to genuine gw bursts or environmental transients are indistinguishable.

(c) *Discrimination algorithms* that separate gw-like from spurious events by fitting the detector outputs to a plane wave with the distinctive properties of the phase and amplitude of the Riemann tensor in vacuum, i.e. (i) the wavefront travels at the speed of light [28]; (ii) the independent components are transverse and traceless (TT). An effective discrimination can be achieved in a worldwide network of non-co-located and nonparallel detectors by exploiting both the phase (i) and amplitude (ii) informations [29,30]. As a by-product of the fit of gw-like events, we will get the spectral content of the gw burst and, eventually, other plane wave parameters such as direction, polarization, and speed of propagation [31].

(d) *Background estimate algorithms* that evaluate the rate of gw-like signals. The standard approach consists in the same algorithms of item (c) applied to the detector outputs shifted by an unphysical time delay (e.g. $\gg 42$ ms for ground based detectors). Such an algorithm establishes if "anything," by chance, is causing a gw-like signal in the detectors and it gives also an estimate of their rate.

Without the additional confidence from discrimination algorithm (c) and background estimate algorithm (d), it would be impossible to assess the false alarm probability of gw-like events. To maximize the detection probability in a

network of gw detectors, the tuning of the complete data analysis pipeline should be made at the step (d), where we can estimate the "true" background, while tuning at the step (b) would be misled, due to the presence of local disturbances.

In what follows we do not enter into details of discrimination algorithms but we will focus on detection algorithms of step (b) that should be able to identify, in the most efficient and thorough way, "pieces" of the observed data with a statistics that differs from the noise statistics. In a "coherent" analysis of a network of detectors, detection (b) and discrimination (c) algorithms can be implemented at the same time with the aim of maximizing the detection efficiency. For instance, an "aperture synthesis" can be formed by a linear combination of detector outputs, shifted in time and weighted in amplitude [32], to create a virtual channel where the detection algorithm is applied. To this purpose, we are developing a "coherent network analysis," based on the wavelet packet transform of the detector outputs that have been previously combined in a multi-dimensional channel [18].

B. Detection algorithms for gw burst searches

In the field of gw searches, many detection algorithms have been proposed during the past decades. Detection algorithms can be discussed in the framework of classical "hypothesis testing" [19,33]: on the basis of the observed data, we have to take our decision whether to reject or fail to reject the null hypothesis (the signal is absent) against the alternative hypothesis (the signal is present) by means of a detection threshold. The alternative hypothesis for the detection of gw bursts is composite, and so is difficult to propose a unique criterion of "optimality." Here we examine some reference detection algorithms (matched filter and power filter) and we devote Sec. IV to the wavelet packet detection algorithm. Throughout the following, the input data of a detection algorithm consists of the samples $x_k = s_k + n_k$ $(k = 1 \ldots N_s)$, where the signal s_k is the argument of the hypothesis test and the noise n_k is a realization of a white, stationary, and Gaussian stochastic process with zero mean and variance σ^2. However, the requirement of a white noise can be easily fulfilled if the detector output can be modeled as a regular process [34]; in this case, we can assume that the samples x_k have been first whitened by a suitable linear filter.

1. Matched filter detection algorithm (MF)

Suppose that the signal $s(t)$ is sampled at discrete time instants $t_k = kt_s$ and that it can be written as $s_k = Af(t_k)$, where $f(t_k) \equiv f_k$ is a known waveform and $A > 0$ its unknown amplitude. The matched filter (also known as Wiener-Kolmogorov filter) is a linear estimator designed to minimize the mean-square error between the detector samples x_k and the estimation $\hat{A}f_k$ of the signal. The well-known solution for white noise is $\hat{A}(\boldsymbol{x}) = \sum_k \boldsymbol{x}_k f_k / (\sum_k f_k^2)$,

where $\hat{A}$ is a Gaussian random variable with mean A and standard deviation $\sigma_{\hat{A}} \equiv \sigma / (\sum_k f_k^2)^{1/2}$; $\sigma_{\hat{A}}$ also represents the Cramèr-Rao lower bound [35] of any linear estimator of A. The decision rule of the MF detection algorithm puts a threshold T on $\hat{A}$, and $R_T = \{x: \hat{A}(x) \geq T\}$ is the resulting rejection region [33]. On the other hand, the likelihood ratio can be written as $\Lambda(A) = \exp[\hat{A}^2 / (2\sigma_{\hat{A}}^2)]$ and therefore we have the equivalence between the decision rule of the MF algorithm and the rejection of the null hypothesis by a threshold crossing of Λ. As the alternative hypothesis is an one-sided composite hypothesis, the Neyman-Pearson lemma holds, and we can conclude that the MF efficiency ϵ is greater than any other test with the same false alarm probability P_{FA}. As $\hat{A}$ is Gaussian distributed, we get

$$\begin{cases} P_{\mathrm{FA}}(T) = \mathrm{erfc}(T / \sqrt{2\sigma_{\hat{A}}^2}) \\ \epsilon(T) = 1/2 \, \mathrm{erfc}((T - A) / \sqrt{2\sigma_{\hat{A}}^2}), \end{cases} \tag{4}$$

where $\mathrm{erfc}(x)$ is the complementary error function. The receiver operation characteristic (ROC) curve [36] can be expressed in a closed form by eliminating the variable T from Eq. (4). The MF algorithm clearly implements an "ideal" receiver of poor practical utility in the field of gw detection as it depends crucially on *a priori* knowledge of the waveform to be detected. Nonetheless, the MF method is interesting as each point of the corresponding ROC curve can be interpreted as an upper bound of the efficiency achievable by any other detection algorithm at a given P_{FA}.

2. Impulsive filter detection algorithm (IF)

The whitened impulsive response of a narrow band detector, roughly a superposition of high-Q dumped sinusoids, can be used as a signal template [11]. The resulting filter, that we call in brief the impulsive filter (IF), depends on detector rather than signal characteristics. The IF detection algorithm, which is based on a threshold on the output of the impulsive filter, is particularly interesting for resonant detectors. The motivations for using IF as an ETG are a simple implementation and an overall good detection efficiency for δ-like gw bursts. Up to now, the IF detection algorithm is used for the search of coincidences among the gw detectors of the IGEC collaboration [8].

3. Power filter detection algorithm (PF)

The power of the data set x_k can be used to form a locally optimal statistics for the detection of transient signals [19]. In fact, by means of power filters we are able to perform a uniformly most powerful test on the localized power defined as

$$\mathcal{E}_q = \sum_{k=1}^{N_s} (q \circ \boldsymbol{x})_k^2, \tag{5}$$

where $\circ$ is the convolution operator and q is a suitable linear phase FIR filter that selects the box of the time-frequency plane containing (most of) the signal power. The test clearly requires the prior knowledge of time of arrival t_0, the duration Δ_τ, central frequency ν_0, and frequency band Δ_ν of the signal.

As shown in Ref. [12], the statistics $\mathcal{E}_q$ is optimal in the sense that it gives the highest probability of correctly detecting a transient signal for given false alarm probability. In the limit of weak signals, the likelihood ratio Λ can be expanded in a Taylor series about $\mathcal{E}_s = 0$, where $\mathcal{E}_s \equiv \sum s_k^2$ is the signal energy. The first nonzero term of this expansion reads $\Lambda_0(q) = \exp(\mathcal{E}_q)$ plus other terms independent of x_k. Of course, this expansion must be restricted to the box of the time-frequency plane where the power is localized. When the signal bandwidth is greater than the detector band (as typical for narrow band detectors), we have to substitute the signal band with the few Hz of the detection band. In the opposite case of wideband detectors, we can form a bank of filters $q(t_0, \Delta_\tau, \nu_0, \Delta_s)$ to find the best localization of the power in the time-frequency plane [12]. Notice that the input data must have different probability distributions (composite alternative hypothesis) but the hypothesis on the local power $\mathcal{E}_q$ is composite unilateral and therefore the Neyman-Pearson lemma still holds. Thus, the "optimal statistics" to detect a signal with known location in the time-frequency plane is simply the local power. In each window the local power $\mathcal{E}_q$ obeys the noncentral χ^2 distribution with noncentral parameter equal to $\mathcal{E}_s/\sigma^2 > 0$,

$$p_\chi(\mathcal{E}_q; \mathcal{E}_s) = \frac{1}{2\sigma^2} \exp\left(-\frac{\mathcal{E}_q + \mathcal{E}_s}{2\sigma^2}\right)$$
$$\times \left(\frac{\mathcal{E}_q}{\mathcal{E}_s}\right)^{(N_s-2)/4} I_{N_s/2-1}(\sqrt{\mathcal{E}_q \mathcal{E}_s/\sigma^2}), \quad (6)$$

where $I_k(x)$ are the modified Bessel functions of the first kind of order k. However, to relax the requirement of the *a priori* knowledge of q, we can form a nonoptimal statistics by choosing the domain of q as large as possible, both in time and frequency, i.e. $\mathcal{E} \equiv \sum_k^{N_s} x_k^2$. The statistics $\mathcal{E}$ is no longer optimal but it depends only on the detector characteristics; in fact, the detection efficiency is independent of signal waveform as long as the signal duration is limited to the data set x_k. We refer to the detection algorithm based on a detection threshold on the total power $\mathcal{E}$, as the power filter (PF) algorithm. For the PF algorithm, the false alarm probability and detection efficiency as a function of the detection threshold T reads

$$\begin{cases} P_{\text{FA}}(T) = \gamma(N_s/2, T)/\Gamma(N_s/2) \\ \epsilon(T) = \int_T^\infty p_\chi(\mathcal{E}, \mathcal{E}_s)d\mathcal{E}, \end{cases} \quad (7)$$

where $\gamma(x, y)$ and $\Gamma(x)$ are the incomplete and complete Euler gamma functions, respectively. The ROC curves of the PF detection algorithms play the role of a lower bound.

4. Other detection algorithms

The whitened noise should have a correlation function that vanishes away from zero and so the autocorrelation of these data can be used to form some statistics to detect the presence of signals (which changes locally the autocorrelation function). A family of filters, based on the autocorrelation or linear interpolation of data, have been intensively studied in Refs. [16,37,38]. The family includes the norm filter, mean, and norm of autocorrelation [37]. The alternative linear fit (ALF), slope (SF), and offset (OF) (intensively studied by the VIRGO collaboration) are based upon fitting a straight line to the data [38].

The BlockNormal detection algorithm [39] is an ETG that identifies moments in time where the data statistics changes. In particular, BlockNormal characterizes the time series between change points by the mean and variance of the samples. The data set (block) identified by the BlockNormal algorithm is characterized by a mean, a variance, a frequency band, a start time, and a duration. Triggers from different bands are merged if they overlap in time into a single trigger [39].

The burst analysis method [14] is based on wavelet packet transform and differs from the WP algorithm in the detection statistics. In fact, the definition of the burst analysis statistic entails 4 steps: (i) decompose the observed data on wavelet packet bases at different levels; (ii) fix the fraction of coefficients that makes the cluster core(s) at any decomposition level; (iii) aggregate the cluster cores in larger clusters; (iv) form a supercluster by looking for clusters that overlap both in time and frequency in adjacent levels. Finally, the detection statistics is the energy content (sum of squared coefficients) of a supercluster. Each step of the burst analysis detection algorithm is controlled by parameters that affect both the false alarm rate and the detection efficiency.

The WaveBurst ETG [15] is an implementation of the burst analysis method: WaveBurst is applied simultaneously for two or more gw detectors looking for coincident coefficient clusters and attains the simultaneous implementation of a detection and discrimination algorithm in a network of gw detectors.

IV. WAVELET PACKET DETECTION ALGORITHM

First, the data set x_k is decomposed over wavelet packet bases associated to complete binary trees of depth $j = 1, 2, 3 \ldots j_{\max}$, where $j_{\max}$ is fixed by the maximum frequency resolution $t_s^{-1}/2^{j_{\max}+1}$ we are interested in. Regularity conditions (e.g. band limited signals) lead us to consider a decomposition based on a smooth mother wavelet like the order 10 symlets (sym10). However, higher and lower orders have been tested without significative changes in the performances of the algorithm. The coefficients of the WP transform d_{j+1}^{2p} and d_{j+1}^{2p+1} at the level $j + 1$ are computed from the coefficients d_j^p of the

previous level by the following recursive relations:

$$d_{j+1}^{2p}[k] = \sum_l h_0[l - 2k]d_j^p[l]$$

$$d_{j+1}^{2p+1}[k] = \sum_l h_1[l - 2k]d_j^p[l], \tag{8}$$

where h_0 and h_1 are defined in Sec. II and $d_0^0[k] = x_k$. As high value coefficients reveal the presence of a signal in the corresponding time-frequency tiles, the decision rule of WP detection algorithm is: reject the null hypothesis if

$$\max\{|d_j^p[k]|\} > T, \tag{9}$$

where $0 \leq j \leq j_{\max}$, $0 \leq p \leq 2^j - 1$, $1 \leq k \leq N_s/2^j$, and the maximum is calculated over all the $p = 0, 1, 2 \ldots 2^j - 1$ leaves and all the $j = 0, 1, 2 \ldots j_{\max}$ binary trees. We note that the detection algorithm has the computational complexity $O(N_s \log N_s)$ of a wavelet packet transform associated to a complete binary tree of order $j_{\max}$ plus the computational complexity $O(N_s)$ of the sorting algorithm.

The N_s coefficients of the WP transform, associated with a *complete* binary tree, ensure a uniform partition of the time-frequency plane. On the other hand, taking into account trees of different depth $\leq j_{\max}$ give rise to multiple resolutions in the time-frequency plane and so wider classes of signals can be detected at the cost of a redundancy in the signal representation. As the total number of coefficients analyzed is $N_s \times (j_{\max} + 1)$, our redundant partition of time-frequency plane translates in a slight increase of false alarms: this is the price to be paid to the robustness of the algorithm.

Each coefficient d_j^p is Gaussian distributed and, in the absence of signals, the mean is zero, but there is a small correlation among coefficients at the same level [20].

The detection statistics of the WP algorithm focuses attention on the extreme values of the data, i.e. the tails of the probability distribution. The crucial theorem on extreme value distributions states that, in the limit as the number of samples tends to infinity, the induced distribution on the maxima of the samples can only take one of three forms: Gumbel, Weibull, or Frechet [40]. The false alarm probability distribution relative to the WP detection statistics is bounded from above by the Gumbel distribution

$$p_G(x) = \frac{1}{\lambda} e^{(x-\mu)/\lambda} \exp(-e^{(x-\mu)/\lambda}), \tag{10}$$

which is appropriate for maxima of $N_s \times (j_{\max} + 1) \gg 1$ normally distributed variates; here μ and λ are the location and scale parameters, respectively. One can easily show that $E = \mu + \gamma_e \lambda$ and $V = \pi^2 \lambda^2/6$, where $\gamma_e = 0.577\,21\ldots$ is the Euler-Mascheroni constant, and E and V are the mean and variance of the maximum of wavelet packet coefficients, respectively; hence μ and λ can be readily estimated by Monte Carlo simulations. As an ex-
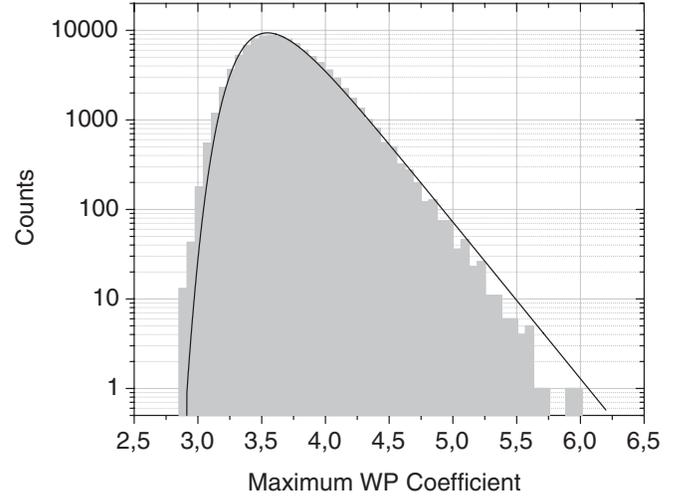
FIG. 1. Distribution of false alarms generated by the WP detection algorithm applied to white noise with zero mean and unit variance. The number of trials is 10^5. The sample mean and variance are 3.689 and 0.101, respectively. The continuous line shows a Gumbel distribution with $\lambda = 0.25$ and $\mu = 3.55$.

ample, we report in Fig. 1 the histogram of $\max\{|d_j^p[k]|\}$ obtained with $N_s = 384$ samples, $j_{\max} = 6$ and white noise; the Gumbel cumulative distribution with $\lambda = 0.25$ and $\mu = 3.55$ is clearly a limiting curve from above of the false alarm probability of the WP method.

As discussed in the next section, the detection efficiency of the WP method is a complicated function of waveforms, amplitude, arrival time, decomposition levels, etc., and so it must be evaluated using dedicated Monte Carlo simulations. However, in the high SNR regime, we can argue that the position in the time-frequency plane of $\max\{|d_j^p[k]|\}$ remains fixed. Thus, the statistical distribution of the maximum value tends to a Gaussian with mean $\mathcal{M}$ and variance σ^2, where $\mathcal{M}$ is the maximum component of the projection of the given signal on the wavelet packet bases. The asymptotic behavior of the ROC curves of the WP method is then determined by the equations

$$\begin{cases} P_{\mathrm{FA}}(T) \simeq \exp(-e^{(T-\mu)/\lambda}) \\ \epsilon(T) \simeq 1/2\,\mathrm{erfc}((T - \mathcal{M})/\sqrt{2\sigma^2}). \end{cases} \tag{11}$$

Equations (11) make it clear why the variances of false alarms and true alarms tend to be different as the SNR increases.

V. DETECTION ALGORITHMS APPLIED TO AURIGA DATA

In order to run the detection algorithms on the data acquired by the AURIGA gw detector, we first consider the data h_k representing the equivalent gw signal at the detector input. The h_k are obtained by the deconvolution of output data from the AURIGA transfer function $H(\nu)$, which accounts for all filtering stages applied to the gw

signal, including electronics and analog to digital converters [11]. The deconvolution is carried out in the discrete time domain $\mathbb{Z}(t_s)$, where $t_s^{-1} = \nu_s = 4882.8125$ Hz is the adopted sampling frequency. By definition, $1/H(z)$ converts the detector output samples in the dimensionless amplitudes h_k of space-time strain; for the AURIGA detector, the procedure to evaluate $1/H(z)$ (calibration) is described in Ref. [41]. The AURIGA intrinsic noise at the detector input is well described by a quasistationary process; to be more precise, the time scales of noise changes are much greater than the correlation times of the stochastic process that are fixed by the fluctuation-dissipation theorem. The noise power spectrum $S_h(\nu)$ in equivalent gw strain, as it was normally measured during the second scientific run of AURIGA [42], is shown in Fig. 2.

The AURIGA conditioning algorithms include a band selection filter (an antialiasing bandpass Butterworth filter of order 12, with cutoff frequencies ~830 and ~990 Hz), followed by a decimation of a factor 12; the resulting sampling frequency is $\nu_d = 406.901$ Hz. After the decimation process, we can analyze the decimated band 0–203.45 Hz without information losses because this band largely includes the aliased detection band, say 850 ÷ 950 Hz.

The two minima at ~865 Hz and ~933 Hz correspond to the AURIGA maximum sensitivity $S_{h,\max}^{1/2}(\nu) \sim 2 \times 10^{-21}$ (Hz)$^{-1/2}$. The full bandwidth at the half maximum is ~25 Hz around the two minima but the whole detection band, say ~100 Hz, can be used in practice for gw burst searches, in fact, the sensitivity within the detection band is in the range $2 \times 10^{-21} < S_h^{1/2}(\nu) < 5 \times 10^{-21}$ (Hz)$^{-1/2}$. It can be shown that the process h_k is regular and that its spectrum can be factorized as $S_h(z) = L(z)L(1/z)$, where $L(z)$ is a minimum phase filter [11]. Then the whitening digital filter $\Gamma(z) \equiv L^{-1}(z)$ can be applied to the process h_k
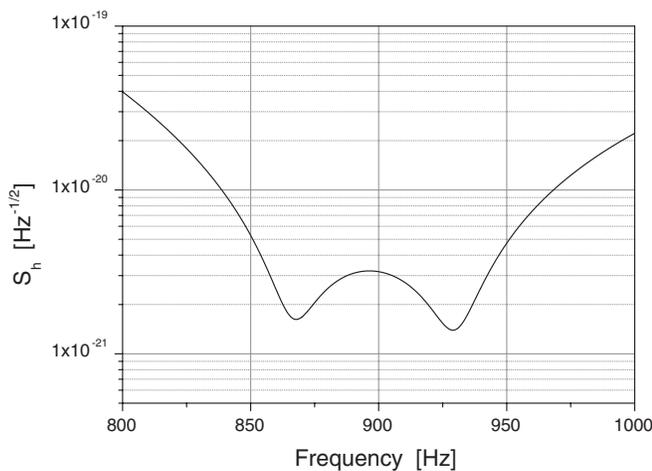
FIG. 2. Sensitivity curve of the AURIGA detectors during its second scientific run obtained by a best fit of the typical noise spectral density.

to calculate its "innovations," i.e. the white process x_k that constitutes the input for the four detection algorithms.

A. Monte Carlo simulations

The difficulties in the comparison of different detection algorithms reflect the composite nature of the alternative hypothesis: the probability density function is not unique, and often direct (analytic) integration over the rejection region R_T—that give P_{FA} and ϵ—is not accessible. An alternative technique to estimate the probability density functions and to perform their integration is the Monte Carlo simulation: for additive, Gaussian, quasistationary stochastic processes, the evaluation of P_{FA} and ϵ can be substituted by their empirical observations. In our specific problem, the Monte Carlo simulations of different detection methods consist in the injection of a large number of waveforms in the simulated data of the AURIGA detector with different SNR. To save computing time in running Monte Carlo, we preferred to simulate directly the white noise n_k after the decimation process. The variance of the simulated noise is assumed to be unitary. The decimation factor from the original sampling frequency ν_s is 15 and so the decimated band amounts to $0 \div 162.76$ Hz. From here on, the variance of the simulated noise is assumed to be unitary and the simulated gw signals s_k are assumed to be whitened, bandpassed, and decimated before their addition to the noise.

The detection procedure is very simple: we consider windows of finite length of $N_s = 320$ samples (i.e. 0.98 s) and apply to each window the binary classifier associated to the decision rules of the MF, IF, PF, and WP detection algorithms. We used the predictions of Eqs. (4) and (7) to check the consistency of Monte Carlo outcomes for the MF and PF methods. The good agreement we found makes us confident about the Monte Carlo implementation. The false alarms, being independent of signals, have been determined with a good accuracy at any interesting rate. On the other hand, the detection efficiency depends both on detection algorithms and on injected waveforms. To test the efficiency of the 4 detection algorithms we decided to inject maximum entropy chirplets, which play the role of reference waveforms. Chirplets are a class of nonstationary signals defined in a six-parameter space,

$$\mathcal{C}(t) = A \exp\left[-\frac{(t - t_0)^2}{4\Delta_\tau^2}\right] \cos\left[\frac{\beta}{2}(t - t_0)^2 \right.$$
$$\left. + \omega_0(t - t_0) + \phi_0\right], \tag{12}$$

where $\beta \equiv (\Delta_\omega/\Delta_\tau)\sqrt{1 - (2\Delta_\tau\Delta_\omega)^{-2}}$ is the chirp rate, ϕ_0 is the initial phase, and A, t_0, Δ_τ, ω_0, and Δ_ω have the usual meaning as in Sec. III B. Chirplets have quite a different signature in the time-frequency plane and a maximum entropy [43]. Other remarkable properties of the chirplet

waveforms are: (i) covariance to scale changes of Δ_ω, chirp rate β, time of arrival t_0, and central frequency ω_0 (very desirable property in the implementation of the Monte Carlo); (ii) satisfy the uncertainty principle with equality when $\Delta_\omega \Delta_\tau = 1/2$; (iii) attain a good resolution in both time and frequency; (iv) "sin-Gaussian" waveforms (widely used in gw burst searches [9]) can be recovered as the particular case $\beta = 0$. In our Monte Carlo simulations, the injected chirplets have been chosen by spanning the parameters space as follows: (1) central frequency $\omega_0/(2\pi) = 815 + 10j$ Hz, where $j = 0, 1, 2, \ldots 15$ (linear spacing); (2) bandwidth $\Delta_\omega/(2\pi) = (1, 2, 5) \times 10^k$ Hz, and duration $\Delta_\tau = (2, 5, 10) \times 10^{-(k+2)}$ s, where $k = 0, 1, 2$ (quasilogarithmic spacing), as shown is Fig. 3; (3) initial phase ϕ_0 uniformly distributed in $[0, 2\pi)$, in order to avoid synchronization effects between the sampling process and detection algorithms.

The different chirplets are injected at the same SNR = $(\sum_k s_k^2)^{1/2}/\sigma$, i.e. at the same ratio of the ℓ_2 norm of the whitened signal and the standard deviation of the white noise, corresponding to the maximum SNR achievable by the ideal matched filter. We generated 10^4 independent noisy data sets for each chirplet and then we applied the detection algorithms.

A few general considerations apply across the MF, IF, EP, and WP detection algorithms and all the injected chirplets: (i) the detection efficiencies improve as the
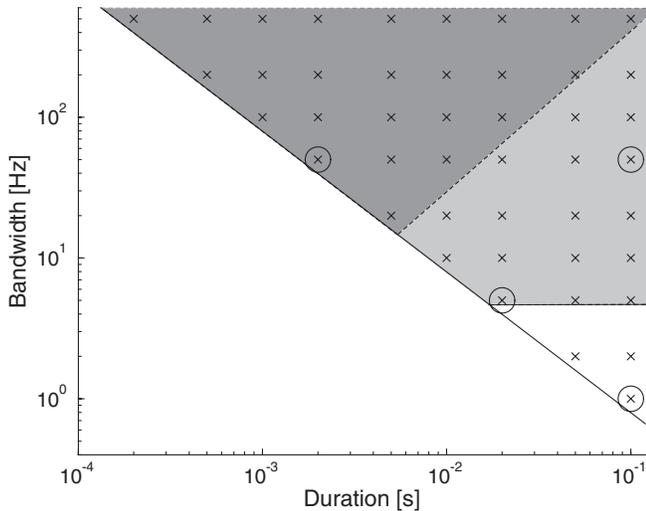
FIG. 3. The boundaries of the parameter space for the chirplets used in the Monte Carlo simulations. The empty region under the continuous line is forbidden by the Heisemberg uncertainty principle. The allowed region can be roughly divided up into 3 parts: (i) the dark gray upper triangle, where the whitened chirplet signals are completely indistinguishable from a $\delta(t)$ signal; (ii) the pale gray quadrilateral, where the chirplet bandwidth is modified by the AURIGA bandwidth; and (iii) the lower triangle, where the chirplets are left unchanged by the AURIGA whitening filter. The open circles indicate the parameters used for the calculation of detection efficiency in Fig. 4.

SNR increases, showing a unitary efficiency in the limit of SNR $\gg 1$; (ii) as the PF algorithm is insensitive to the injected waveforms, the corresponding ROC is almost the same. (iii) The ROC curves of WP detection algorithm are above the corresponding EP curves, except for long duration signals $\Delta_\tau > 2 \times 10^{-2}$; (iv) the IF is below the EP ROC curves when the whitened waveforms are quite different from the whitened impulsive response (i.e. for signals in the lower white triangle of Fig. 3).

As a representative example, we report in Fig. 4 some ROC curves we have obtained by injecting SNR = 8 chirplets, with the intent of figuring out the relative performances of the four algorithms.

A further look on the differences among MF, EP, and IF methods can be obtained by the injection of 10^4, SNR = 8, $\omega_0/(2\pi) = 900$ Hz and $\beta = 0$ chirplets (sin-Gaussians), with the aim of calculating the detection efficiencies at 10^{-2} s^{-1} false alarm rate, as a function of signal duration Δ_τ [44]. In Fig. 5, we observe the sharp decrease of the IF efficiency as soon as $\Delta_\tau > 5 \times 10^{-3}$ s^{-1}, i.e. when the bandwidth of whitened signal becomes smaller than the detector bandwidth (pale gray quadrilateral of Fig. 3). In this range, due to the adopted maximum depth of the wavelet packet trees ($j_{max} = 6$), also the WP efficiency slightly changes. However, the small decrease observed in Fig. 5 can be reduced by considering wavelet transforms with $j_{max} > 6$.

By spanning the whole chirplet manifold, we have been able to evaluate the spread of detection efficiency of WP, IF, and PF algorithms: in Fig. 6 we report the average ϵ_{ave}, maximum ϵ_{max}, and minimum ϵ_{min} values of detection efficiency at 10^{-2} s^{-1} false alarm rate, that we found by running the complete Monte Carlo (10^4 trials per each injected signal), as a function of SNR. As we are not assuming a uniform distribution over the chirplet parameters manifold, ϵ_{ave} cannot be interpreted, using Bayes' formula, as the mean detection probability of chirplets. The quasilogarithmic spacing of Δ_τ and Δ_ω favors indeed the particular hypothesis of sin-Gaussian signals with duration $\Delta_\tau < 2 \times 10^{-3}$ s or bandwidth $\Delta_\omega/(2\pi) < 10$ Hz. It is worth noticing that the IF detection algorithm completely misses some signals, as clearly indicated by the minimum detection efficiency equal to the false alarm probability $\epsilon_{min} = 10^{-2}$. The PF efficiency is almost independent of injected chirplets as shown by the narrow dark gray region in Fig. 6.

In principle, the problem of the detection of "rare" gw bursts (low rate) buried in the detector noise (low SNR) requires the minimization of the maximum probability of a loss, i.e. ϵ_{min}. The intermediate variability of WP method is the result of a tradeoff between the high efficiency and low variability requirements. In this respect, the performances of the WP detection algorithm are satisfactory.
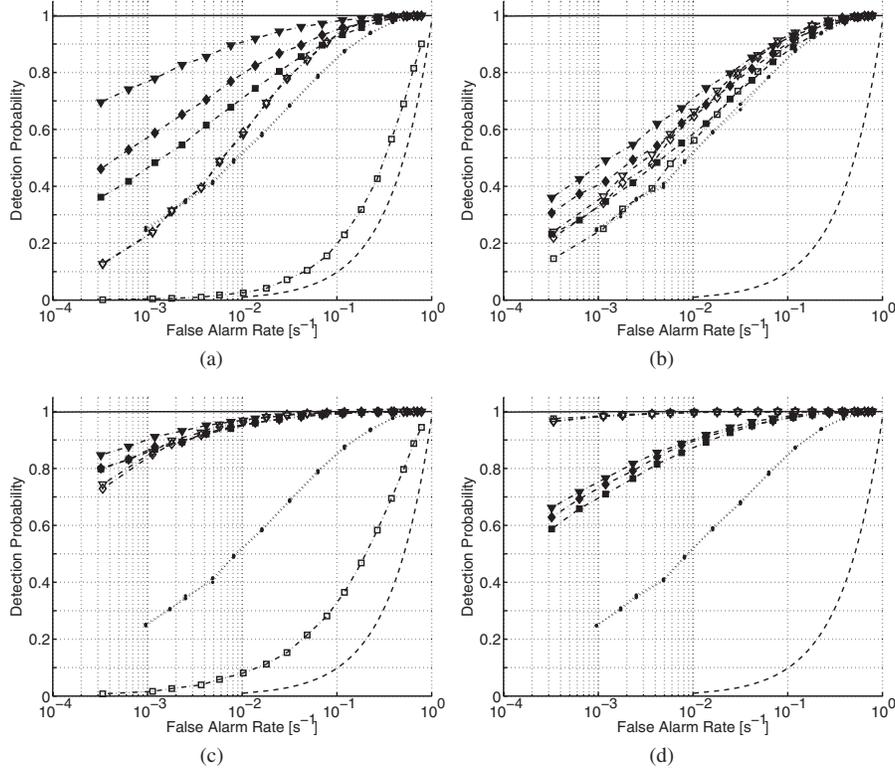
FIG. 4. ROC curves for the IF (open symbols), PF (dots), and WP (filled symbols) detection algorithms obtained by injecting chirplets with SNR = 8 and central frequencies 865 Hz (lower triangle), 900 Hz (square), and 933 Hz (upper triangle). The continuous and dashed lines represent the ROC curves of the ideal MF "receiver" and the random decision rule, respectively. The chirplet duration and frequency band correspond to the four circled points of parameter space in Fig. 3, namely, (a) $\Delta_\tau = 0.1$ s and $\Delta_\omega = 1.0$ Hz; (b) $\Delta_\tau = 0.1$ s and $\Delta_\nu = 50$ Hz; (c) $\Delta_\tau = 0.02$ s and $\Delta_\omega = 5$ Hz; and (d) $\Delta_\tau = 0.002$ s and $\Delta_\omega = 50$ Hz.
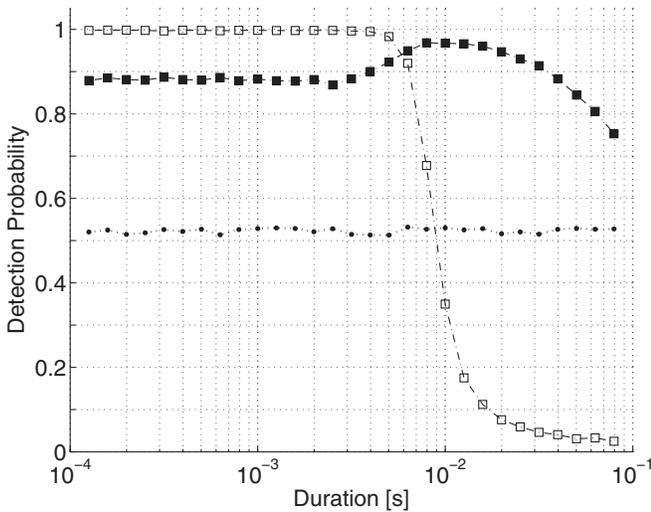
FIG. 5. The dependence of the detection efficiency of IF (open squares), WP (filled squares), and PF (dots) methods on signal duration. The injected signals are SNR = 8 sin-Gaussians and the false alarm rate has been fixed at the level of 10^{-2} s^{-1}.
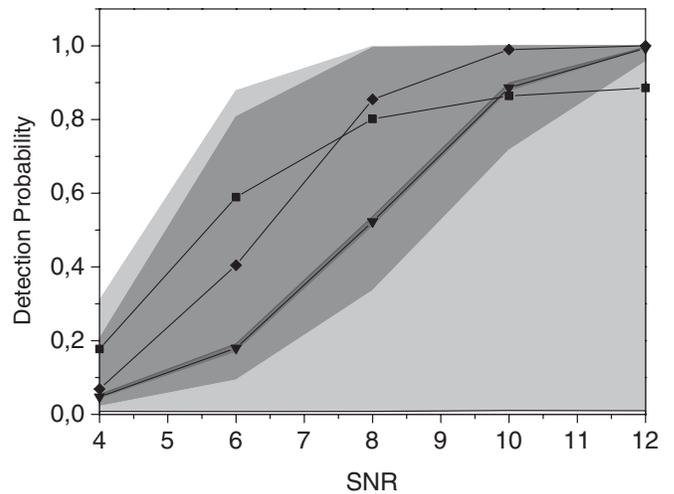
FIG. 6. The detection efficiencies of the IF, WP, and PF detection algorithms as a function of SNR. The pale gray, gray, and dark gray regions indicate the maximum variability of IF, WP, and PF detection algorithms. Squares, diamonds, and triangles represent the average detection efficiency of IF, WP, and PF methods, respectively.

B. Application of the WP detection algorithm to AURIGA data

As already stated, the output of any gw detector requires a conditioning procedure in order to ensure noise Gaussianity and signal sparsity. For instance, the AURIGA output is sometimes contaminated by unmodeled disturbances, related to the detector environment or maintenance operations, that affect the AURIGA capability in gw searches. These disturbances last for long periods of time in respect to the gw burst signals. Other disturbances are well localized in the frequency domain as spurious spectral lines within the detection band. The energy content of spurious lines is very high, not stationary and/or does not agree with the predictions of the fluctuation/dissipation theorem. This component of environmental disturbances, that give rise to "large fluctuations" in the whitened detector data x_k, can be ruled out by notch filters tuned to the spurious line frequencies [45]. The long lasting disturbances can be removed by applying the "epoch vetoes," which take care of the periods of time affected by the appearance of a lot of unmodeled excitations. It should be noted that the epoch vetoes reduce the live time of AURIGA with no impact on the detection efficiency because their possible effects on the statistics of the WP algorithm can be easily dealt with suitable boundary conditions for the wavelet packet transform. On the contrary, the application of the notch filters in the detection band makes some differences between simulated noise and the real data of the AURIGA detector. In fact, the coefficients of the wavelet packet transform that represent frequency bands overlapping the notch bands have a smaller variance in respect to the other coefficients and so they do not produce false alarms.

Nevertheless, the false alarm rate is only slightly changed in respect to the simulated equivalent noise as the decision rule of the WP algorithm depends on the largest coefficients (i.e. the largest fluctuations in the time-frequency plane). The main impact of the notch filters is on the detection efficiency of those signals with significative overlapping among their frequency band Δ_ν and the notch frequencies. The loss in the detection efficiency caused by conditioning procedures is unavoidable but it can be minimized by identifying and defeating the environmental noise sources.

Moreover, the output of the AURIGA detector is affected by short-lived disturbances of electromagnetic origin with a rate of few hundred per hour. They are a sort of glitch in the time domain with a duration of $\sim$40 ms and they are correctly detected by the WP algorithm as isolated events. The intense glitches can be easily identified by means of a tagging algorithm [45] and therefore they can be removed from subsequent analyses.

In order to investigate the relevant effects of "ad hoc" preprocessing procedures on the AURIGA data, we report in Fig. 7 the histograms of the maximum coefficients
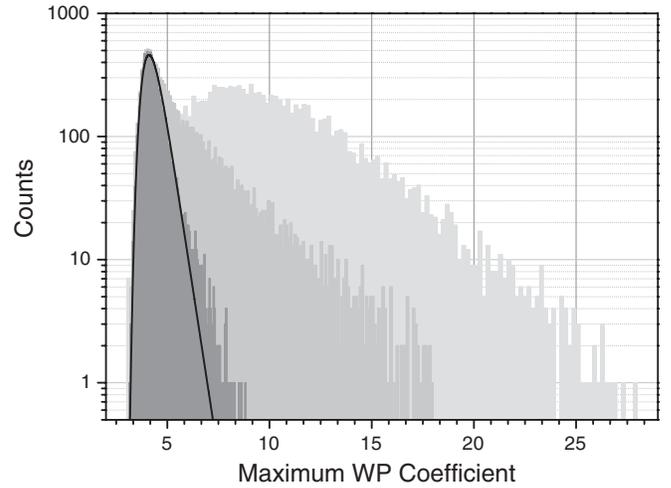
FIG. 7. Histograms of the maximum WP coefficients obtained with a set of $\sim$3 h of AURIGA data. Light gray histogram: original data with no conditioning algorithm; gray histogram: after the application of epoch vetoes and frequency notches; dark gray: as for the gray histogram after the removal of electromagnetic glitches. The continuous line shows the Gumbel distribution with $\lambda = 0.41$ and $\mu = 4.10$.

produced by the WP algorithm. The AURIGA data has been normalized to unitary variance before to feed them to the WP detection algorithm. We note that the distribution of the maximum WP coefficients (dark gray histogram), after the application of epoch vetoes, notch filters, and the removal of electromagnetic glitches, is very close to a Gumbel distribution with $\lambda = 0.41$ and $\mu = 4.10$; these values does not differ very much from $\lambda = 0.25$ and $\mu = 3.55$ of the Monte Carlo simulations (see Fig. 1); however, the motivations behind this discrepancy are the cumulative effects of the very low amplitude glitches, which show an increasing rate as their amplitude decreases. In any case, the signal sparsity, combined with the extreme distribution theorem, allows the estimate of false alarm probability and therefore we were able to discriminate signals from noise at a given probability level. Indeed, the continuous curve in Fig. 7 represents the superposition of maxima arising from Gaussian fluctuations of AURIGA noise and the low amplitude contribution of glitches; moreover, we can argue that the few transient signals exceeding the Gumbel distribution in Fig. 7 are mainly due to glitches not identified and removed by the AURIGA tagging algorithm.

A different approach to conditioning algorithms is the study of the histograms of the wavelet coefficients of real data sets. In fact, the bulk probability distribution of each coefficient must be a Gaussian and we can use Gaussianity tests, e.g. a threshold on skewness and kurtosis indexes [11], to discard those wavelet packet coefficients which refer to spurious frequency lines and/or epoch vetoes. For instance, we find that some histograms of the wavelet packet coefficients of the AURIGA data are clearly non-Gaussian distributed with fluctuations that are orders of

magnitude greater than the other coefficients. Of course, these coefficients, which correspond to spurious frequencies, have to be removed from the detection procedure otherwise they would dominate the detection statistics, as clearly shown in Fig. 7. The outcome of the Gaussianity criterion is a sort of ''mask'' to be applied to the wavelet packet coefficients which discard the non-Gaussian coefficients from the detection statistics. Here, the term non-Gaussian refers to the presence of long lasting spurious signals rather than non-Gaussian stochastic processes. The Gaussianity of the coefficients can also be verified by a Gaussian fit on the corresponding histograms. A drawback of this approach is the coarse-grained resolution of wavelet packet coefficients; in fact, the outer nodes of the wavelet packet trees do not allow a better determination of the time-frequency boxes containing spurious disturbances in respect to the rectangles defined by the epoch vetoes and notch filters. In our opinion, the vetoing procedures for gw detectors using data relative to the detection bandwidth are still an open problem, and it could only be solved in the framework of a network analysis.

VI. CONCLUSIONS

We have implemented and thoroughly studied a novel method for the detection of gw bursts of ~ 1 s maximum duration. The method is based on wavelet packet transform of observed data sets and does not require *a priori* hypothesis on waveform patterns, arrival time, or spectral content. The WP detection algorithm is quite general and can be applied to narrow or wideband detectors once the requirements of noise Gaussianity and signal sparsity are fulfilled. Both these conditions should be met by gw detector outputs after the application of conditioning algorithms. We made also clear the separation between detection and estimate phases for burst signals: in our approach the false alarm rate is controlled by a suitable detection threshold on the maximum WP coefficient. In general, the estimate of waveform characteristic is not required and should be deferred to subsequent discrimination algorithms. To calculate the detection efficiency of the MF, WP, IF, and PF algorithms, we have explored the 6-dimensional chirplet parameter space.

The resulting performances of WP algorithm, measured by the ROC curves, are quite satisfactory. The detection efficiency is mainly affected by the maximum level of wavelet decomposition, which fixes the smallest frequency structure that the WP algorithm can recognize. In addition, we observed a slight decrease of the detection efficiency when the injected chirplets are completely spread over the time-frequency plane, i.e. in the high Δ_τ and Δ_ω region.

The IF detection algorithm shows a large variability of its efficiency across the chirplet parameter space, due to the mismatches between the injected signals and the impulsive filter template.

As the PF algorithm uses the energy in the whole band, its detection efficiency becomes quite insensitive to waveforms. However, the achieved robustness with respect to signal variability entails a lowering of detection efficiency. In fact, the PF algorithm represents a sort of lower bound for ROC curves of a detector.

An open problem for a fair comparison of the performances of detection algorithms is, of course, a faithful representation of the gw burst manifold. In this paper, to run our Monte Carlo, we make use of a simple meshing of a portion of the chirplet parameter space; however, astrophysical sources could favor completely different and/or unexplored regions.

The Achilles' heel of any detection algorithm is the inability to discriminate spurious from genuine gw signals with the data of a single detector. In this respect, the importance of the WP detection algorithm lies in the prospects it opens for the near future, when different kinds of detectors, with heterogeneous bandwidths and antenna patterns, will be operating together in a worldwide network. In fact, if gw signals have no characteristic waveforms, an ''aperture synthesis'' between different detectors is the only way to reject spurious signals and, at the same time, to gain information on gw sources. We think that the absence of any *a priori* assumptions on waveform patterns is the ''conditio sine qua non'' to set up network detection algorithms for genuine gw bursts. Much work has still to be devoted to extend detection algorithms to worldwide networks of gw detectors because of the partial overlapping of both detection bandwidths and antenna patterns (e.g. among resonant bars and interferometers). Anyway, we think that only the projections of the TT Riemann tensor on different detectors (amplitude informations), combined with the timing of detectors located at different sites (phase informations), have to be used as intrinsic signatures of a gw burst. The problem of really taking advantage of these intrinsic signatures, without significative losses of detection efficiency, would deserve further investigation and it will be the topic of a forthcoming paper [18].

[1] C. Cutler and K. S. Thorne, in *"Proceedings of GR16", Durban, South Africa*, edited by N. T. Bishop and S. D. Maharaj (World Scientific, Singapore, 2001); K. S. Thorne, in *Black Holes and Relativistic Stars*, edited by R. M. Wald (University of Chicago Press, Chicago, 1998), p. 41.

[2] M. H. P. M. van Putten *et al.*, Phys. Rev. D **69**, 044007 (2004).

[3] J. Baker, M. Campanelli, C. O. Lousto, and R. Takahashi, Phys. Rev. D **69**, 027505 (2004); **65**, 124012 (2002).

[4] T. Damour and A. Vilenkin, Phys. Rev. D **64**, 064008 (2001).

[5] R. F. Stark and T. Piran, Phys. Rev. Lett. **55**, 891 (1985); T. Zwerger and E. Müller, Astron. Astrophys. **320**, 209 (1997); H. Dimmelmeier, J. A. Font, and E. Müller, Astrophys. J. Lett. **560**, L163 (2001); C. D. Ott, A. Burrows, E. Livne, and R. Walder, Astrophys. J. **600**, 834 (2004).

[6] See, e.g., E. Cappellaro, R. Evans, and M. Turatto, Astron. Astrophys. **351**, 459 (1999), and references therein.

[7] V. Fafone, Classical Quantum Gravity **21**, S377 (2004); B. Abbot *et al.*, Classical Quantum Gravity **23**, S29 (2006).

[8] Z. A. Allen *et al.*, Phys. Rev. Lett. **85**, 5046 (2000); P. Astone *et al.*, Phys. Rev. D **68**, 022001 (2003).

[9] B. Abbot *et al.*, Classical Quantum Gravity **23**, S29 (2006); Phys. Rev. D **72**, 062001 (2005).

[10] K. S. Thorne, in *300 Years of Gravitation*, edited by S. W. Hawking and W. Israel (Cambridge University Press, Cambridge, England, 1987), p. 330458.

[11] A. Ortolan *et al.*, Classical Quantum Gravity **19**, 1457 (2002).

[12] W. G. Anderson *et al.*, Phys. Rev. D **63**, 042003 (2001); Int. J. Mod. Phys. D **9**, 303 (2000).

[13] S. Mohanty, Phys. Rev. D **61**, 122002 (2000); J. Sylvestre, Phys. Rev. D **66**, 102004 (2002).

[14] S. Klimenko and G. Mitselmakher, Classical Quantum Gravity **21**, S1819 (2004).

[15] S. Klimenko *et al.*, Classical Quantum Gravity **21**, S1685 (2004).

[16] N. Arnaud *et al.*, Phys. Rev. D **67**, 062004 (2003).

[17] For a wavelet based denoising with automatic noise level estimation, see e.g. D. L. Donoho, IEEE Trans. Inf. Theory **41**, 613 (1995).

[18] A. Ortolan and M. Camarda (unpublished).

[19] C. W. Helstrom, *Statistical Theory of Signal Detection* (Pergamon Press, Oxford, 1968).

[20] S. Mallat, *A Wavelet Tour of Signal Processing* (Academic Press, San Diego, 1998).

[21] C. S. Burrus, R. A. Gopinath, and H. Guo, *Introduction to Wavelets and Wavelet Transforms, a Primer* (Prentice Hall, Upper Saddle River, NJ, 1998).

[22] S. Mallat, Trans. Am. Math. Soc. **315**, 69 (1989); Y. Meyer, *Ondelettes et Opérateurs Ondelettes* (Hermann, Paris, 1990); S. Mallat, IEEE Trans. Acoust. Speech Signal Process. **37**, 2091 (1989).

[23] I. Daubechies, *Ten Lectures on Wavelets* (Society for Industrial and Applied Mathematics, Philadelphia, 1992).

[24] R. Coifmann, Y. S. Q. Meyer, and M. V. Wickerhauser, *Signal Processing with Wavelet Packets* (Yale University, Numerical Algorithm Research Group, New Haven, CT, 1990).

[25] See, e.g., H. B. Callen and T. A. Welton, Phys. Rev. **83**, 34 (1951); J. Weber, Phys. Rev. **101**, 1620 (1956).

[26] See, e.g., Cutler and Thorne, in "Proceedings of GR16", Durban, South Africa (Ref. [1]).

[27] P. Antonioli *et al.*, New J. Phys. **6**, 114 (2004).

[28] The simplest discrimination algorithm is the search for time coincidences among triggers of two or more parallel detectors; see e.g., P. Astone, S. DAntonio, and G. Pizzella, Classical Quantum Gravity **19**, 1443 (2002), and references therein.

[29] V. C. Visconti *et al.*, Phys. Rev. D **57**, 2045 (1998).

[30] I. S. Heng, A. Ortolan, and F. Salemi, Classical Quantum Gravity **20**, S617 (2003).

[31] S. V. Dhurandhar and M. Tinto, Mon. Not. R. Astron. Soc. **234**, 663 (1988); M. Tinto and S. V. Dhurandhar, Mon. Not. R. Astron. Soc. **236**, 621 (1989); Y. Gürsel and M. Tinto, Phys. Rev. D **40**, 3884 (1989).

[32] L. S. Finn, Phys. Rev. D **63**, 102001 (2001).

[33] M. H. A. Davis, in *Proceeding of the NATO Advanced Research Workshop on Gravitational Wave Data Analysis*, edited by B. F. Schutz (Kluwer Academic Publishers, Dordrecht/Boston/London, 1989), pp. 73–94.

[34] A. Papoulis, *Probability, Random Variables, and Stochastic Processes.* (McGraw-Hill, Singapore, 1984).

[35] H. Cramér, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, NJ, 1946).

[36] T. Fawcett, HP Laboratories, Palo Alto, CA, Technical Report No. HPL-2003-4.

[37] N. Arnaud *et al.*, Phys. Rev. D **59**, 082002 (1999).

[38] T. Pradier *et al.*, Phys. Rev. D **63**, 042002 (2001); T. Pradier *et al.*, Int. J. Mod. Phys. D **9**, 309 (2000).

[39] J. W. C. McNabb *et al.*, Classical Quantum Gravity **21**, S1705 (2004).

[40] See e.g. N. Balakrishnan and A. C. Cohen, *Order Statistics and Inference* (Academic Press, New York, 1991).

[41] J. P. Zendri *et al.* (to be published).

[42] L. Baggio *et al.*, Phys. Rev. Lett. **94**, 241101 (2005).

[43] S. Mann and S. Haykin, IEEE Transactions on Signal Processing **43**, 2745 (1995).

[44] The sin-Gaussian with $\Delta_\tau = 0.02$ s was used as a reference signal to calculate the detection efficiency in the LIGO-AURIGA joint burst searches: see L. Cadonati *et al.*, Classical Quantum Gravity **22**, 1 (2005).

[45] L. Baggio *et al.*, in Annual Report 2004, p. 121, ISBN 88-7337-008-X.