Filtering coalescing binary signals: Issues concerning narrow banding, thresholds, and optimal sampling

S. V. Dhurandhar* and B. F. Schutz
Department of Physics and Astronomy, University of Wales, College Cardiff, Post Box 913, Cardiff CF2 3YB, United Kingdom
(Received 9 February 1994)

When the raw output of a gravitational wave detector is correlated with the matched filter of a coalescing binary wave form the filtered output shows a periodic behavior—it rings at a certain frequency. This phenomenon could be worrisome since the signal peak in the filtered output might be reduced if it falls in the "trough" of the sinusoid. In this paper we address this question and present a detailed examination of the ringing which is caused by the effective narrow banding by the matched filter of detector noise. We first solve the problem for an idealized "box" filter and show that the ringing frequency is roughly the central frequency of the box if the box is not too wide. For an idealized coalescing binary filter we show that the expected value of this frequency is 1.27 f_s, where f_s is the seismic noise cutoff of the detector. The ringing implies that there is some redundancy in the filtered ouput. Also the autocorrelation function of the filtered output resembles the sinc function, and hence adjacent sample points are correlated, i.e., the filtered output is colored. These two phenomena are related and have a bearing on the setting up of thresholds and also suggest that we resample the filtered output at a coarser rate. We investigate the problem of thresholds when the filtered output is colored and obtain relations between the false alarm probabilities and threshold levels. Finally we suggest optimal sampling rates so that the resampled filtered output is uncorrelated.

PACS number(s): 04.80.Nn, 04.30.Db, 06.50.Dc, 97.80.−d

I. INTRODUCTION

The detection of gravitational waves has been an outstanding problem in experimental physics for over three decades now. Starting from the pioneering experiments of Weber using a bar detector [1], there has been a lot of effort in building detectors of higher sensitivity (see [2] for a 1987 review of the bar detector program). In recent years, several groups around the globe have been planning to build long base line laser interferometric gravitational wave detectors [3–6], prototypes of which already exist in Germany, Great Britain, and USA. Because of their inherent broadband nature, interferometric detectors can be used to follow the changing frequency of the gravitational waves emitted during the in spiral of compact binary systems during the final stages of their evolution. The rate of such coalescences is estimated to be about three per year out to a distance of 100–200 Mpc [7].

Because of their extragalactic origin the amplitude of these signals is, in general, expected to be too low for them to be seen directly in the time series. However, these sources are relatively simple systems, so the nature of the gravitational wave emitted during the inspiral can be predicted with fair reliability. The wave form is called a chirp for obvious reasons. With the wave form given, it is possible to extract the signal out of the detector noise by the use of a data analysis technique called matched filtering. This technique consists of correlating the output of a detector with a filter which, in the Fourier domain and when the noise is stationary, is nothing but the Fourier transform of the expected signal inversely weighted by the noise power spectral density. The noise power spectral density rises sharply at low frequencies and could be virtually taken to be infinite below a certain frequency f_s called the seismic cutoff, below which the noise rises like a steep wall. Therefore, the detector noise provides the lower band limit to the matched filter. Also the coalescing binary signal decays as $f^{-7/6}$ providing an effective upper band limit to the matched filter. Thus the matched filter effectively narrow bands the output. When the parameters of the filter exactly match with those of the signal we have what is called a matched filter. It is well known in the theory of hypothesis testing that of all the linear filters a matched filter performs the best in extracting a given signal buried in noisy data [8].

In this paper we address two related problems which occur because the filter effectively narrow bands the output. They are as follows.

(a) *The ringing of the filtered output.* When the data are passed through a matched filter of the coalescing binary signal, the filtered output $c(t)$, where t is the time lag, shows a nearly sinusoidal behavior [9]. It looks like a relatively high-amplitude sinusoid superposed with random noise. In the first part of this paper we explain the ringing with the help of a geometrical analysis of a simple narrow band filter. The geometrical formalism is set up in Sec. II while in Sec. III we deal with the ringing problem. In the first part of Sec. III we examine the case of a box filter—a filter which is bandwidth limited but having the same magnitude for all its Fourier components. This case is easier to deal with and affords invaluable insight into the problem. In the second part of Sec. III we consider the case of the chirp filter.

*Permanent address: Inter-University Centre for Astronomy and Astrophysics, Post Bag 4, Ganeshkind, Pune 411 007, India.

(b) *The filtered output is colored.* Schutz [10] has shown that the autocorrelation function $a(\tau) \equiv \langle c(t)c(t+\tau)\rangle$ of a narrow band filter resembles a sinc function with a characteristic damping time which he calls the "decorrelation time." Samples separated by time lags greater than the decorrelation time are expected to be uncorrelated. This observation raises two issues: one concerning false alarm probabilities and thresholds and the other is what we term as *optimal sampling.* So far treatments on coalescing binary detection problems have for the sake of simplicity ignored the covariances in the filtered output [9,11] for setting up thresholds for detection. However, as Schutz has pointed out there are covariances in the filtered output and the problem should be faced squarely. Finn [12] has suggested using the joint probability density function for the sampled output which is a natural generalization. But this ignores correlations between successive points. The filtered output is narrow banded with a sharp lower cutoff f_s and a smooth but relatively rapid upper cutoff decaying as $f^{-7/6}$.

The narrow-banded nature of the filter implies that the filtered output is not only correlated in the statistical sense but is actually algebraically (in fact linearly) dependent. This can be readily understood in terms of a "box" filter of $N > 2k$ points containing k positive (and symmetrically placed about zero frequency, k negative) Fourier components of height unity and the rest of the $N-2k$ components zero. If this filter is applied to a data segment containing N points, the filtered output consists of N points out of which only $2k$ of them are linearly independent. Thus the joint probability density function for the sampled filtered output is $2k$ dimensional. We may either choose any of the $2k$ out of the N filtered output points as linearly independent sets, or we may formulate this problem by going over to the Fourier space and treat the $2k$ Fourier components of the box as our linearly independent set of variables We choose the latter option. Our probability density function is defined on the subspace of the Fourier space corresponding to the bandwidth of the filter. Also, we can associate the number of degrees of freedom for the points in the filtered output and in the case of the box filter defined above it is just $2k$. For the matched filter of the coalescing binary with a sharp lower cutoff and a smoothly decaying upper cutoff we are able to define an effective box filter. The number of Fourier components in the effective box are the number of degrees of freedom in the filtered output and the probability density function can be defined over the Fourier components constituting the effective box. For the matched filter considered, the number of degrees of freedom of the filtered output is less than the number of points in the filtered output.

Since the effective number of degrees of freedom is less than the number of points in the sampled filtered output it should be possible to resample the filtered output more coarsely and still retain the relevant information. In fact, the Nyquist theorem [8] essentially pertains to this issue. The colored filtered output signifies redundant information and it should be possible to sample the output more coarsely without loss of information pertaining to detection. Finding an optimal sampling rate has implications for storage space and subsequent processing speed. In Sec. IV we address these two problems and compute thresholds as functions of false alarm probabilities and finally propose an optimal sampling rate. As in Sec. III we first examine the problem for the box filter before attacking the problem for the coalescing binary filter.

II. NOISE, PROBABILITY DISTRIBUTIONS, AND FALSE ALARM PROBABILITIES

In this section we prepare the ground by setting up a geometrical framework which helps us in understanding and investigating the problems of ringing and thresholds. To appreciate the ringing phenomenon the geometrical viewpoint is especially helpful. We subdivide this section into three subsections: in Sec. II A we set up the geometrical framework, in Sec. II B we discuss noise and in Sec. III C we review the relevant material on false alarm probabilities.

A. Geometrical framework

We adopt the definitions, notations, and conventions of Ref. [13] for the Fourier transform, the discrete Fourier transform (DFT), and the inverse Fourier transform. We have used the program of Ref. [13] to perform simulations which we will come to later in the text.

Consider a raw output data segment $o(t)$, $0 \le t \le T$ where the data segment lasts T seconds and consists of noise and perhaps a signal. Assuming a constant sampling rate with sampling interval Δ, the number of points N in the data segment is given by $N = T/\Delta$. The sampled output is the real time series $o_k = o(k\Delta)$, $k = 0, 1, \ldots, N-1$. We now look upon o_k as components of a vector $\mathbf{o} = o_k \mathbf{e}_k$, where $\mathbf{e}_k$ are unit vectors with components δ_{kn}, $n = 0, 1, \ldots, N-1$. They form an orthonormal basis for the sampled outputs which span the N-dimensional Euclidean space $\mathcal{R}^N$. Thus, $\mathbf{o} \in \mathcal{R}^N$. We have used the summation convention of summation over repeated indices. This will be so unless otherwise indicated.

But since our analysis is going to involve Fourier transforms we may regard $\mathbf{o}$ also as an element of $\mathcal{C}^N$, the N-dimensional complex space, in which the imaginary part of each o_k is zero. We assume the usual scalar product on $\mathcal{C}^N$: namely, for any $\mathbf{z}, \mathbf{w} \in \mathcal{C}^N$,

$$\mathbf{z} \cdot \mathbf{w} = z_k w_k^* , \tag{2.1}$$

where the asterisk denotes the operation of complex conjugation. For z_k, w_k real the scalar product reduces to the usual Euclidean one on $\mathcal{R}^N$.

A Fourier transform can be regarded as a change in basis of $\mathcal{C}^N$; we transform between the bases $\mathbf{e}_k$ to $\mathbf{f}_n$ by the relations

$$\mathbf{f}_n = \alpha_{nk} \mathbf{e}_k , \quad \mathbf{e}_k = \beta_{kn} \mathbf{f}_n , \tag{2.2a}$$

$$\alpha_{nk} = 1/N \exp(-2\pi i n k/N) ,$$
$$\beta_{kn} = \exp(2\pi i n k/N) , \tag{2.2b}$$

where $0 \le k \le N-1$, $-N/2 \le n \le N/2-1$. These trans-

formations connect the time domain components of o to its Fourier ones. We regard the vector o to be unchanged while its components undergo a transformation under a change of basis, i.e., this is a passive transformation. We write $\mathbf{o}=o_k\mathbf{e}_k=\tilde{O}_n\mathbf{f}_n$, then the components transform as

$$\tilde{O}_n=\beta_{nk}o_k \ , \ , \tag{2.3a}$$

$$o_k=\alpha_{nk}\tilde{O}_n \ . \tag{2.3b}$$

The transformations defined above are for the DFT instead of the Fourier transform since we find that convenient to use in the analysis that follows. The Fourier-transform components are related to the DFT components by $\tilde{o}_n\simeq\tilde{O}_n\Delta$.

We note that α_{nk} and β_{nk} are symmetric complex matrices which are inverses of each other. Therefore the order of indices in the above formulas is irrelevant. We state the following properties for the matrices which we will use later:

$$\alpha_{nk}\beta_{km}=\delta_{nm} \ ,$$

$$N\alpha_{nk}\alpha^*_{km}=\frac{1}{N}\beta_{nk}\beta^*_{km}=\delta_{nm} \ , \tag{2.4}$$

$$\beta_{n,l+p}=\beta_{nl}\beta_{np} \ .$$

The above relations show that $N^{1/2}\alpha_{nk}$ and $N^{-1/2}\beta_{nk}$ are unitary matrices. The unitarity property is just another statement of Parseval's theorem. From Eqs. (2.2) the orthogonality properties of the basis vectors may be deduced. Thus,

$$\mathbf{e}_k\cdot\mathbf{e}_l=\delta_{kl} \ , \quad \mathbf{f}_m\cdot\mathbf{f}^*_n=\frac{1}{N}\delta_{mn} \ . \tag{2.5}$$

For a given data segment, o is a fixed vector. We assume additive noise which means that o is a sum of the noise vector n and the signal vector s:

$$\mathbf{o}=\mathbf{n}+\mathbf{s} \ . \tag{2.6}$$

Thus o is comprised of two parts; the noise which is a random vector and the signal which for the coalescing binary case is deterministic. (This may not be the case in general, for example, the stochastic background of gravitational radiation due to cosmic strings.) The noise n is distributed according to some probability distribution function. The signal, on the other hand, may depend on several parameters. In this paper the signal is that of the coalescing binary for which we make the simplest possible assumptions. We assume the Newtonian wave form with two point masses spiralling around each other in a circular orbit, emitting gravitational radiation according to the quadrupole formula. With these assumptions the signal then depends on two parameters. We discuss the wave form in the next section.

In the matched filtering technique, the statistic is the correlation c between the raw output and the matched filter. In the limit of continuous sampling, for the filter $q(t)$, the correlation is given by

$$c(\tau)\equiv\int_{-\infty}^{\infty}o(t)q(t+\tau)dt$$
$$=\int_{-\infty}^{\infty}\tilde{o}(f)\tilde{q}^*(f)e^{2\pi if\tau}df \ , \tag{2.7}$$

where τ is the time lag which has been introduced to adjust for the arbitrary arrival time of the signal. For a signal which arrives at $t=t_a$ the correlation will have a peak at $\tau=-t_a$. The correlation can be calculated with the help of the fast Fourier transform (FFT) with relatively few operations.

In the discrete picture equation (2.7) goes over to

$$C_m=\frac{1}{N}\sum_{n=-N/2}^{N/2-1}\tilde{O}_n\tilde{Q}^*_n\exp(2\pi imn/N)$$
$$\equiv\sum_{n=-N/2}^{N/2-1}\tilde{O}_n\tilde{Q}^*_n\alpha^*_{mn} \ . \tag{2.8}$$

Here $\tilde{Q}_n$ is the DFT of $q(t)$. The above sum can be interpreted as follows: For each time lag $m\Delta$ we calculate the scalar product between the vector o and the time translated filter vector $\mathbf{q}_m\equiv N\tilde{Q}_n\alpha_{mn}\mathbf{f}_n$. The family $\mathbf{q}_m$, $m=0,1,\ldots,N-1$ represents the filter translated by various instants $m\Delta$. Thus, Eq. (2.8) can be compactly written as

$$C_m=\mathbf{o}\cdot\mathbf{q}_m \ , \tag{2.9a}$$

with $\mathbf{q}_m$ given by

$$\mathbf{q}_m=N\tilde{Q}_n\alpha_{mn}\mathbf{f}_n=\tilde{Q}_n\exp[-2\pi imn/N]\mathbf{f}_n \ . \tag{2.9b}$$

B. Noise and matched filters

In the frequency range of 100 Hz to a few kHz the detector noise is dominated by the photon counting noise. If we restrict ourselves to this range of frequencies then we can make the following assumptions about the character of the noise. (A detailed discussion of the noise in the present prototypes can be found in [10].)

(i) At each instant of time t_k, the noise $n_k\equiv n(t_k)$ is a random variable with zero mean; i.e., $\langle n_k\rangle=0$. The angular brackets denote an ensemble average.

(ii) In general, the noise could be colored:

$$\langle n_k n_l\rangle=\phi_{kl} \ . \tag{2.10a}$$

If $\phi_{kl}=\text{const}\times\delta_{kl}$ then the noise is said to be white.

(iii) The noise is stationary. This implies that ϕ_{kl} depends only on $|k-l|=p$ say, i.e., $\phi_{k,k+p}\equiv K_p$, where now the array K_p contains information about the color of the noise. The so-called one-sided power spectral density usually denoted by $S_h(f)$ is then up to a numerical factor just the DFT $\tilde{K}_n$ of K_p. In terms of the DFT of n_k, we have

$$\langle\tilde{N}_n\tilde{N}^*_m\rangle=N\tilde{K}_n\delta_{mn} \ . \tag{2.10b}$$

(iv) The seismic vibrations cause the noise in an interferometric detector to rise steeply below a certain frequency f_s. We thus assume that $S_h(f)=\infty$ for $f\le f_s$. In the present prototype detectors f_s is between 200 and 400 Hz. In future interferometric detectors it is expected to be around 40 Hz initially and in advanced detectors it will be lowered to about 10 Hz by using special seismic isolation techniques [3,14]. We choose $f_s=100$ Hz whenever numerical computations about the chirp are in-

volved, but we retain f_s as a free parameter in much of our discussion. We define $n_s=[f_sT]$ where the square brackets denote the integer part of the quantity within the brackets. Therefore, $\tilde{K}_n=\infty$, $n<n_s$. Also we assume that the power spectral density is constant for frequencies above f_s; i.e., we set $\tilde{K}_n=\sigma^2$, $n\geq n_s$. This assumption is not unjustified if the power spectrum is more or less flat above the seismic cutoff. However, for the chirp, the filter decays as $f^{-7/6}$ for $f\geq f_s$ effectively cutting off the higher frequencies. Therefore, if the power spectral density is flat within the effective bandwidth of the filter the white noise assumption is approximately valid.

(v) Finally, we assume that the noise can be described by a Gaussian distribution function. If we assume this property in the time domain, since the Fourier transform only involves a linear operation, the Fourier components of the noise also follow a Gaussian distribution.

The matched filter q is given by the equation

$$\phi_{ik}q_k=s_i\ , \tag{2.11a}$$

where $\mathbf{s}=s_k\mathbf{e}_k=\tilde{S}_n\mathbf{f}_n$ is the signal. This matrix equation can be solved for q to yield

$$q_k=\mu_{ik}s_i\ , \tag{2.11b}$$

where μ_{ik} is the inverse of the matrix ϕ_{ik}. When the noise is stationary Eq. (2.11b) is a convolution and it can be solved with the help of Fourier transforms. Thus,

$$\tilde{Q}_n=\frac{\tilde{S}_n}{\tilde{K}_n}\ . \tag{2.12}$$

For convenience we choose $\sigma=1$. The matched filter is then just the signal chopped off at frequencies less than f_s. Thus,

$$\tilde{Q}_n=\begin{cases}\tilde{S}_n\ , & |n|\geq n_s\ ,\\ 0\ , & |n|<n_s\ .\end{cases} \tag{2.13}$$

With this filter, the expression for the correlation reduces to

$$C_m=\frac{1}{N}\sum_{n_s}^{N/2-1}\tilde{O}_n\tilde{Q}_n^*\exp(2\pi imn/N)+\text{c.c.}\ , \tag{2.14}$$

where c.c. denotes the complex conjugate. The variance of C_m can be computed more easily if the output contains noise exclusively. Since we have assumed the mean of the noise to be zero, $\langle\tilde{O}_n\rangle=\langle\tilde{N}_n\rangle=0$. Thus, C_m which is a linear function of $\tilde{N}_n$ has mean zero. The variance of each C_m is then given by

$$\begin{aligned}\langle C_m^2\rangle&=\frac{2}{N^2}\sum_n\sum_p\langle\tilde{N}_n\tilde{N}_p^*\rangle\tilde{Q}_n^*\tilde{Q}_p\\ &\qquad\times\exp(2\pi imn/N-2\pi imp/N)\ ,\\ &=\frac{2}{N}\sum_{n_s}^{N/2-1}|\tilde{Q}_n|^2\ .\end{aligned} \tag{2.15}$$

The Gaussian nature of the noise allows us to write the joint probability density function for the raw output. If the raw output consists exclusively of noise then the probability density function is given by

$$p_0(\mathbf{o})=\frac{|\det\mu|^{1/2}}{(2\pi)^{N/2}}\exp(-\tfrac{1}{2}\mu_{ik}o_io_k)\ . \tag{2.16}$$

The mean of $\mathbf{o}=0$ when there is no signal present. However, if the raw output contains a signal s then the mean of the distribution gets shifted to s and the corresponding joint probability density function is given by

$$p_1(\mathbf{o})=\frac{|\det\mu|^{1/2}}{(2\pi)^{N/2}}\exp[-\tfrac{1}{2}\mu_{ik}(o_i-s_i)(o_k-s_k)]\ . \tag{2.17}$$

We remark here that the nature of the noise may not remain Gaussian for large values of the standard deviation as the data analysis of the Glasgow prototype detector has shown [15,16]. In that case this analysis will have to be redone.

C. False alarm probabilities

The problem of detection of a signal as distinct from estimation of its parameters consists of determining whether or not a signal is present in the data. This problem is statistical in nature because of the noise which is a random variable. The noise can conspire to appear as a signal and hence noise prevents us from definitely predicting the presence or absence of a signal. Therefore, a claim for detection must be accompanied by a probability that a signal is present in the data and is not just an instance of noise. Because of the stochastic nature of the noise the experimenter is confronted with the problem of testing a statistical hypothesis. He or she must decide on the basis of the output o which of the hypotheses, signal absent (H_0) or signal present (H_1) is true. She or he must adopt a strategy that assigns a definite choice of either H_0 or H_1 based on the output o. This amounts to partitioning the sample space into two disjoint subsets say R_0 and R_1, so that if $\mathbf{o}\in R_i$ we say H_i is true, for $i=0,1$. The experimenter will occasionally err in making a decision but the aim here is to minimize the chance of an error or more precisely to minimize the cost incurred in erring. For details see [8].

When the hypothesis H_1 is extremely rare (as in our case) the principal factor in the cost turns out to be the number of times H_1 is incorrectly chosen in some given number of trials. The ratio of such incorrectly construed detections or *false alarms* to the total number of trials is called the false alarm probability. The relevant criterion here is the Neyman-Pearson one which fixes the false alarm probability the experimenter can afford while maximizing the so-called detection probability, which we define below. In terms of the probability density functions p_0 and p_1 the false alarm probability P_F and the detection probability P_D are given by

$$P_F=\int_{R_1}p_0(\mathbf{o})d^N\mathbf{o}\ , \tag{2.18a}$$

$$P_D=\int_{R_1}p_1(\mathbf{o})d^N\mathbf{o}\ . \tag{2.18b}$$

The Neyman-Pearson criterion leads to the condition on the likelihood ratio:

$$\Lambda(\mathbf{o})=p_1(\mathbf{o})/p_0(\mathbf{o})\ . \tag{2.19}$$

This ratio determines the choice between the two hypotheses. $\Lambda=\Lambda_0$ defines a hypersurface in the sample space $\mathcal{R}^N$, the decision surface which separates R_0 from R_1. Thus if Λ is less than Λ_0 decide on H_0, otherwise H_1. Λ_0 is called the threshold.

From Eqs. (2.16) and (2.17) we have

$$\ln\Lambda(\mathbf{o})=\mu_{ik}s_i o_k-\tfrac{1}{2}\mu_{ik}s_i s_k \ . \tag{2.20}$$

Since $\mu_{ik}s_i s_k$ is just a constant we may use the statistic $\mu_{ik}s_i o_k$ instead of Λ to decide between the hypotheses. But this statistic is just the correlation, the scalar product between the output and the filter [Eq. (2.11b)]. For arbitrary arrival time we must consider the entire family of filters $\mathbf{q}_m$ which yields the filtered output as the statistics C_m for each time lag $m\Delta$.

The problem now reduces to comparing C_m with a threshold η which is determined by the false alarm probability. We should therefore decide on the false alarm probability we can afford and then compute η. The false alarm probability we can afford depends on the expected event rate of signals. For compact coalescing binaries consisting of neutron stars or blackholes, the event rate is expected to be about 3 per year out to 200 Mpc [7]. The Laser Interferometric Gravitational Wave Observatory (LIGO) in its advanced stage could have a range up to 1 Gpc and hence the event rate could be as large as one per day. In this case one false alarm per month of data would probably be acceptable. But for lower sensitivity a false alarm rate of one per year might be required. However, there is little difference in these thresholds for data sampled in the kHz range.

Given a data segment we have to deal with an N-component random vector with components C_m. As mentioned before there are covariances between the various components and hence the problem of fixing the threshold for a given false alarm probability is quite complex. But in our case the problem becomes even more complex since the filtered output is narrow band and spans a subspace of $\mathcal{R}^N$. The joint probability density function is defined over this subspace. It turns out that to obtain false alarm probabilities for given thresholds one must integrate the probability density function over awkward volumes in this subspace. We address this question in Sec. IV.

III. THE RINGING OF THE CORRELATION

A. The chirp wave form and the matched filter

We consider the wave form given by Peters and Mathews [17]. In the transverse traceless (TT) gauge the gravitational wave emitted by a coalescing binary system is described in terms of the two polarizations usually denoted by $h_+(t)$ and $h_\times(t)$. The noise-free response of the detector is a linear combination of the two polarization amplitudes with coefficients depending on the orientation of the detector relative to the direction of propagation of the wave [11,18–20]. The effect of an arbitrary relative orientation of the detector and the plane of the orbit of the binary is only to alter the amplitude and the phase of the signal at the site of the detector without affecting its time dependence. Therefore, for constructing the matched filter, it is enough to consider, say, the + polarization. The wave form from such a system of total mass M and reduced mass μ located at a distance r is given by

$$h(t)\equiv h_+(t)$$
$$=\mathcal{N}a(t)^{-1}\cos\left[2\pi\int_{t_a}^{t}f(t')dt'+\Phi\right] . \tag{3.1}$$

The quantities appearing in the above equation are defined as follows.

t_a and Φ are, respectively, the time of arrival and the phase of the signal when the instantaneous gravitational wave frequency of the signal reaches some fiducial frequency, say f_s.

(1) $\mathcal{M}=(\mu^3M^2)^{1/5}$ is called the chirp mass; the Newtonian wave form depends only on this parameter instead of the two individual masses of the stars.

(2) ξ is the time taken for the two stars to theoretically coalesce starting from a time when the instantaneous frequency is f_s:

$$\xi=3.00\left[\frac{\mathcal{M}}{M\odot}\right]^{-5/3}\left[\frac{f_s}{100\ \text{Hz}}\right]^{-8/3}\ \text{sec} \ . \tag{3.2}$$

The coalescence time ξ serves as a parameter to characterize the wave instead of the chirp mass $\mathcal{M}$.

(3) $a(t)$ is the time-dependent normalized distance between the stars [normalized to $a(t_a)=1$]:

$$a(t)=\left[1-\frac{t-t_a}{\xi}\right]^{1/4} . \tag{3.3}$$

(4) $f(t)$ is the instantaneous gravitational wave frequency given by

$$f(t)=f_s a(t)^{-3/2} . \tag{3.4}$$

(5) $\mathcal{N}$ is a constant.

In specifying a matched filter, in addition to the power spectral density of the noise we need the Fourier transform of the signal. In the stationary phase approximation, the positive frequency components of the Fourier transform of Eq. (2.1) are given by [11,21]

$$\tilde{h}(f)=\int_{-\infty}^{\infty}h(t)\exp(2\pi ift)dt=\mathcal{N}\sqrt{\xi}\tilde{H}(f) , \tag{3.5a}$$

where

$$\tilde{H}(f)=\left[\frac{2}{3f_s}\right]^{1/2}\left[\frac{f}{f_s}\right]^{-7/6}\exp[-i\psi(f)] , \tag{3.5b}$$

$$\psi(f)=-2\pi ft_a+2\pi f_s\xi\alpha(f)+\Phi+\frac{\pi}{4} ,$$

$$\alpha(f)=\frac{1}{5}\left[8-3\left[\frac{f}{f_s}\right]^{-5/3}-5\frac{f}{f_s}\right] .$$

Since $h(t)$ is real, the negative frequency components can be found by using the relation $\tilde{h}(-f)=\tilde{h}^*(f)$. The quantity $\tilde{H}(f)$ is chosen to have unit normalization, i.e.,

$$2\int_{f_s}^{\infty}|\tilde{H}(f)|^2 df = 1 \ . \tag{3.6}$$

The lower limit in the integral is taken to be f_s and not zero since, as discussed earlier, the detector response has a lower-frequency cutoff.

The matched filter is then defined through its Fourier transform as follows:

$$\tilde{q}(f) = \begin{cases} A\tilde{H}(f) \ , & |f| \geq f_s \ , \\ 0 \ , & |f| < f_s \ , \end{cases} \tag{3.7}$$

where A is a normalization constant. This filter narrow bands the filtered output in the low- as well as the high-frequency regimes. At low frequencies the filter is in fact zero. On the other hand, for $|f| \geq f_s$ it decays as $|f|^{-7/6}$. This narrow banding produces rather intriguing and interesting effects when one inverse Fourier transforms into the time domain. The filtered output so obtained rings in a quasiperiodic manner.

B. The ringing of the correlation

We start with a numerical example of ringing which immediately gives a clear idea of the phenomenon. Figure 1 shows an example of ringing. We have taken the raw output to be white noise, correlated it to a chirp filter with parameters $\xi = 3.0$ sec and $\Phi = 0$. The seismic cutoff is fixed at $f_s = 100$ Hz. The data segment consists of $N = 8192$ points sampled at 2.5 kHz. We observe from the figure that for a time lag interval of about 0.5 sec there are about 6 cycles. It is in general true that the filtered output or the correlation rings at about 1.25 times the seismic cutoff frequency and does not depend on the parameters defining the chirp namely ξ or Φ. We observe that the filtered output is not, of course, a perfect sinusoid but appears to be a sinusoid with noise superposed on top of it. The phenomenon is not only intriguing but worrying. For suppose a signal arrived exactly during the trough of the sinusoid it could get "lost" in the trough and go undetected. The other side of the coin is that the signal could ride on top of a crest and thus get "pushed" up. In this way weak signals would have a chance of being detected.

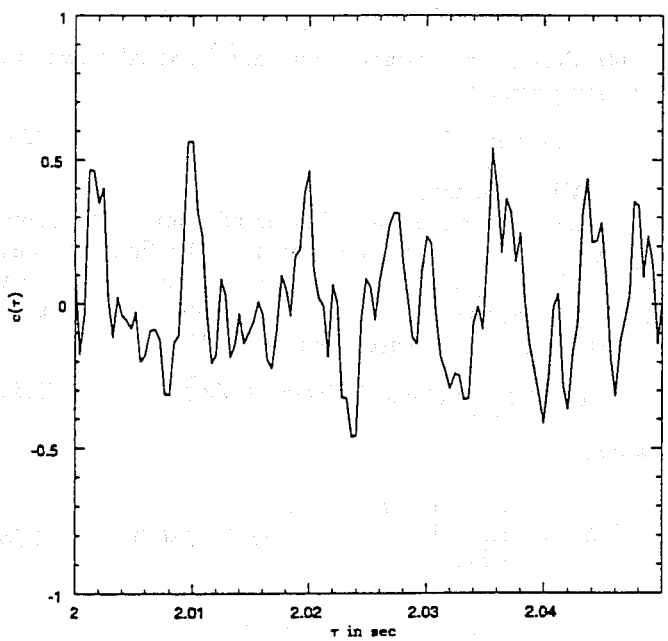

FIG. 1. The figure displays the ringing of the correlation when the chirp filter is correlated with white noise. The correlation c is plotted as a function of time lag τ between 2 and 2.05 sec. The seismic cutoff frequency is taken to be 100 Hz. The correlation shows six or seven cycles which implies a ringing frequency between 120 to 140 Hz.

In the following subsection we first discuss an idealized problem or a toy model which is simpler to analyze in practice but contains the essence of the true problem. We consider the filter when it is a "box."

C. The box filter

The domain of definition is the Fourier space. We define the box filter in terms of its DFT components as

$$\tilde{Q}_n = \begin{cases} 1 \ , & n_0 \leq |n| \leq n_0 + k - 1 \ , \\ 0 \ , & \text{otherwise} \ . \end{cases} \tag{3.8}$$

The procedure to obtain the filtered output is then identical to that of the chirp filter. We assume that the raw output is white noise in the time domain. We then correlate the raw output with the box filter and obtain the filtered output. If the raw output contains in addition a signal, the filtered output of the matched filter will exhibit a peak at the appropriate time lag corresponding to the arrival of the signal. The filtered output will be a superposition of the filtered output of the noise and the peak, the filtered output of the signal. The ringing therefore pertains to the noise and hence in order not to unnecessarily complicate matters we assume that the raw output consists of noise only. Since we want the filtered output to be real, any filter must have the property $\tilde{Q}_{-n} = \tilde{Q}_n^*$. The box filter satisfies this property and hence yields a real filtered output. Thus, actually we have two boxes, mirror images of each other about the zero frequency.

The filtered output for the filter defined by Eq. (3.8) is then

$$C_m = \frac{1}{N}\sum_{n_0}^{n_0+k-1} \tilde{N}_n \exp[2\pi i m n/N] + \text{c.c.} \tag{3.9}$$

Figures 2(a)–2(c) show examples of the filtered output when the box is narrow ($2k \ll N$), of medium width ($2k \sim N/4$) and broadband ($2k \sim N$), respectively. In all these figures C_m is plotted versus m the time lag. As can be observed the ringing behavior is most pronounced in case (a), less pronounced in case (b), and not at all seen in case (c). The ringing when it is apparent is around the central frequency of the box, $n_c = n_0 + k - \frac{1}{2}$.

The geometrical framework set up in Sec. II is now useful for analyzing the problem and providing invaluable insight into the phenomenon of ringing. Equation (2.9b) defines the filter vectors $\mathbf{q}_m$ for the box:

$$\mathbf{q}_m = \sum_{n_0}^{n_0+k-1} \exp[-2\pi i m n/N]\mathbf{f}_n + \text{c.c.} \tag{3.10}$$

We note that $\mathbf{q}_m$ are real vectors. We have also written the α_{mn} explicitly because then it is easy to follow the argument. For a particular data segment the noise vector $\mathbf{n}$ is fixed. However, the filter vectors $\mathbf{q}_m$ depend on m which varies from 0 to $N-1$. The $\mathbf{q}_m$ are *sequentially* called upon to have their scalar product taken with $\mathbf{n}$. Thus it must be that the ringing phenomenon is encoded in the "dynamics" of $\mathbf{q}_m$ in the space $\mathcal{R}^N$, as m increases monotonically from 0 to $N-1$.

We now list below some properties of $\mathbf{q}_m$ which will turn out to be useful for further analysis.

(i) The $\mathbf{q}_m$, $m=0,1,\ldots,N-1$, span a $2k$-dimensional subspace $\mathcal{Q}$ of $\mathcal{R}^N$, since they are linear combinations of the vectors $\mathbf{f}_n$, $\mathbf{f}_n^*$, $n_0 \le n \le n_0+k-1$.

(ii) Any $2k$ different $\mathbf{q}_m$ are linearly independent and constitute a basis of $\mathcal{Q}$. This property follows from the observation that Eq. (3.10) may be viewed as a matrix transformation relating the Fourier basis vectors $\mathbf{f}_n$ to $\mathbf{q}_m$. The matrix in question is just the truncated Fourier matrix α_{mn} (except for a constant factor of N) which is nonsingular. Therefore, $2k$ different $\mathbf{q}_m$ form a basis of $\mathcal{Q}$.

For the box, the filtered output is the projection of the raw output into the subspace $\mathcal{Q}$ and C_m is its projection on the filter vector $\mathbf{q}_m$. Since the filtered output vector is in $\mathcal{Q}$ it is determined by $2k$ scalar products with any $2k$ linearly independent vectors which could as well be $2k$ different $\mathbf{q}_m$'s. This means that any $2k$ different C_m determine the filtered output vector uniquely.

(iii) The scalar products between the $\mathbf{q}_m$ are

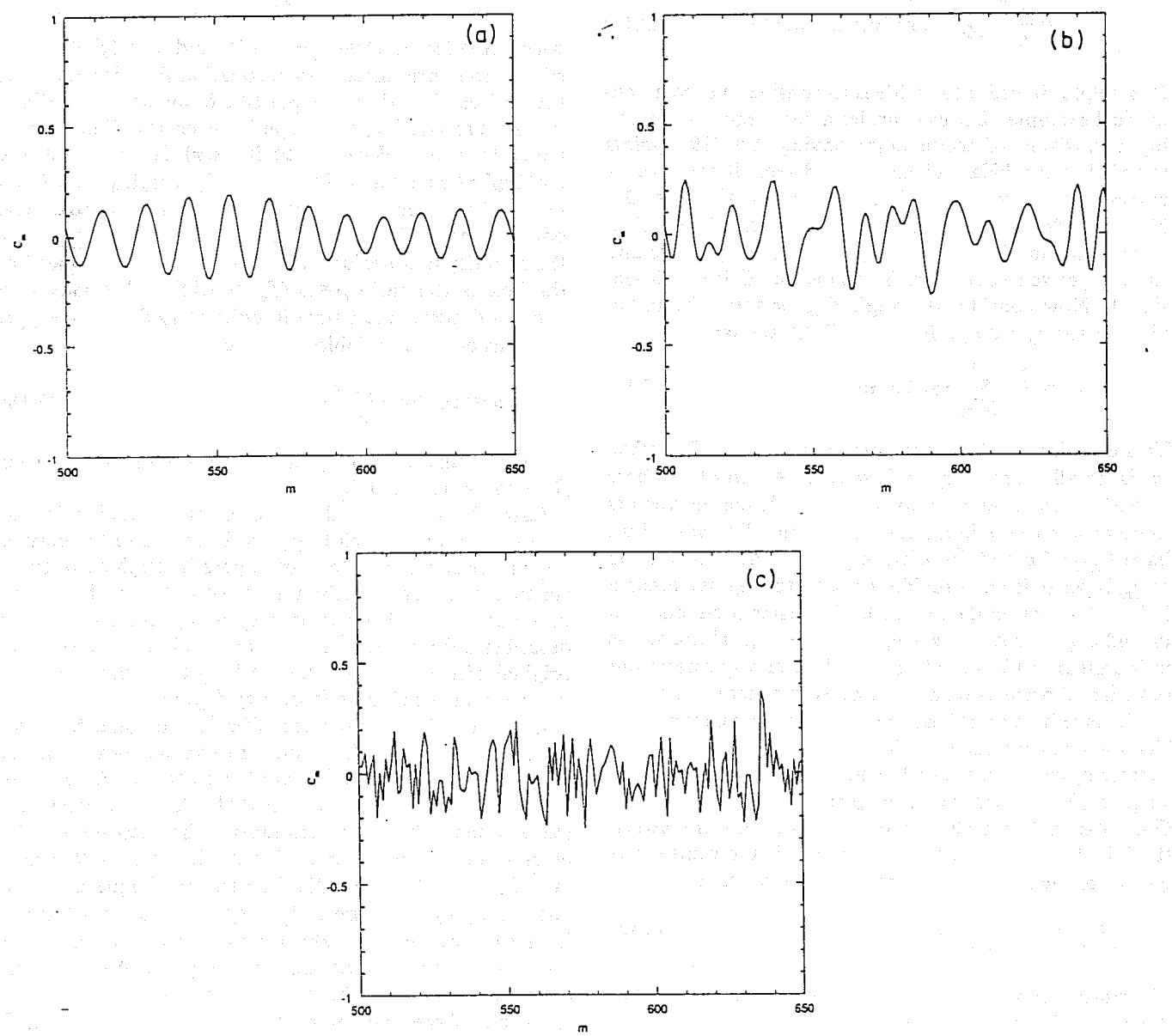

FIG. 2. (a)–(c) depict plots of the ringing of the filtered output for box filters. The parameters are as follows: $N=8192$ and (a) $k=55$, (b) $k=820$, (c) $k=4096$. The ringing frequency for (a) is ~ 575 giving a ringing period of $\Delta m \sim 14$. For (b) the ringing frequency is about 700. (c) is essentially white noise. There is clear ringing seen in (a), not so clear ringing seen in (b), and no ringing at all observed in (c).

$$\mathbf{q}_m\cdot\mathbf{q}_p=\sum_{n=n_0}^{n_0+k-1}\sum_{q=n_0}^{n_0+k-1}\exp[2\pi ipq/N-2\pi imn/N]\mathbf{f}_n\cdot\mathbf{f}_q^*+\text{c.c.}$$

$$=\frac{2}{N}\sum_{n_0}^{n_0+k-1}\cos[2\pi n(p-m)/N]\ . \tag{3.11a}$$

From (3.11a) we obtain the norm of $\mathbf{q}_m$ as

$$|\mathbf{q}_m|=\left[\frac{2k}{N}\right]^{1/2}\ . \tag{3.11b}$$

The angle between the vectors is then

$$\cos\theta_{mp}=\frac{\mathbf{q}_m\cdot\mathbf{q}_p}{|\mathbf{q}_m||\mathbf{q}_p|}$$

$$=\frac{1}{k}\sum_{n_0}^{n_0+k-1}\cos[2\pi n(p-m)/N]\ . \tag{3.12}$$

It is helpful in order to fix ideas to consider a specific numerical example. Let us consider a data segment consisting exclusively of white noise having $N=8192$ points sampled at 2.5 kHz. Then $T\simeq 3.28$ sec. If we take n_0 corresponding to $f_s=100$ Hz then $n_0=[f_sT]=327$. Now consider a very narrow band having $k=3$. We then have a box and its mirror image with three positive and three negative frequencies. The subspace $\mathcal{Q}$ is six dimensional. Now consider the angle θ_{0m} between the initial filter vector $\mathbf{q}_0$ and $\mathbf{q}_m$. From Eq. (3.12) we have

$$\cos\theta_{0m}=\frac{1}{3}\sum_{n=327}^{329}\cos(2\pi mn/N)\ . \tag{3.13}$$

This equation can be approximately solved for θ_{0m} when m is small. Thus, $\theta_{0m}\sim 2\pi mn_c/N=2\pi m/25$, where $n_c=328$ is the average n for the box. Hence the vectors more or less rotate in a plane in a circle. For $m=25$ we have $\theta_{0,25}\sim 2\pi$ and therefore $\mathbf{q}_{25}\sim\mathbf{q}_0$. Thus the motion of $\mathbf{q}_m$ is periodic and provides the clue to ringing. Since $\mathbf{n}$ is fixed for a given data segment, its scalar products with $\mathbf{q}_0$ and $\mathbf{q}_{25}$ are roughly same, i.e., $C_0\sim C_{25}$. However, as m increases to higher values this plane slowly rotates out of its initial orientation and the $\mathbf{q}_m$ scan other directions.

We have here examined the case of a very narrow box. The following properties hold true for a narrow box: The vectors $\mathbf{q}_m$ rotate more or less uniformly in a plane for a single or few rotations and the plane itself slowly rotates. One rotation is completed when the timelag Δm between the initial and final vector is $\sim N/n_c$. If the central frequency of the box is f_c then we have the relations

$$n_c=f_cT\ ,\quad \Delta m=\frac{1}{f_c\Delta}\ . \tag{3.14}$$

The ringing period is $1/f_c$ and the filtered output rings at the central frequency of the box.

When the box consists of many Fourier components ($k\gg 1$), it is advantageous to take the continuous limit of the above discrete analysis. The discrete picture is easier to visualize geometrically while for analysis the continuous one is more convenient. We have the correspondences $(p-m)\Delta\to\tau$, $n/T\to f$, $n_0/T\to f_1$, $n/T\to f_2$, $\theta_{p,p+m}\to\theta(\tau)$. The quantity $\theta(\tau)$ is then the angle between $q(t)$ and $q(t+\tau)$. Let us define $\alpha(\tau)=\cos\theta(\tau)$, then we have the relation

$$\alpha(\tau)\equiv\cos\theta(\tau)=\frac{1}{f_2-f_1}\int_{f_1}^{f_2}\cos 2\pi f\tau\,df$$

$$=\cos(2\pi f_c\tau)\frac{\sin X}{X}\ , \tag{3.15}$$

where $X=2\pi\tau B$, $B=(f_2-f_1)/2$, and $f_c=(f_1+f_2)/2$. $\alpha(\tau)$ is the normalized autocorrelation function and B is the half bandwidth of the positive frequency box. When $\tau\to 0$, $\alpha(\tau)\to 1$, i.e., the angle between the filter vectors tends to zero. When $f_c\gg B$, $\alpha(\tau)$ is then just the $\cos(2\pi f_c\tau)$ modulated by the slowly varying sinc function and we recover the results of the narrow box. The box is broadband, i.e., of full band width when $B=f_c$. So the quantity that decides the narrowness (or broadness) of the box is the ratio $\beta=B/f_c$, $0\le\beta\le 1$. We may also define a dimensionless time in units of $1/f_c$, i.e., $u=f_c\tau$. In terms of these variables we have

$$\alpha(u)=\cos 2\pi u\frac{\sin 2\pi\beta u}{2\pi\beta u}\ . \tag{3.16}$$

We now consider two regimes when β is small and when β is comparable to unity.

Case (i). $\beta\ll 1$. In this case the sinc function in Eq. (3.16) varies on a much larger time scale than the cosine. As u increases from 0 to 1, $\alpha(u)$ goes through a full cycle starting from one maximum at $u=0$ to another maximum at $u\sim 1$. Geometrically speaking, the filter vector describes a full rotation and at $u\sim 1$ points roughly in its original direction. Thus the period of ringing is ~ 1 in units of u or the ringing frequency $f_{\text{ring}}\sim f_c$.

Case (ii). $\beta\sim 1$. Now the effect of the sinc function cannot be neglected. The ringing frequency now depends on β. Let the first maximum of $\alpha(u)$ for $u>0$ occur at $u=u_m(\beta)$. From case (i) $u_m(0)=1$. As β increases, u_m starts reducing steadily from unity. The decrease is rather slow in the beginning. For $\beta\sim 0.2$, $u_m\sim 0.985$ and even for $\beta\sim 0.4$, $u_m\sim 0.92$. The ringing frequency is related to u_m by the formula $f_{\text{ring}}=f_c/u_m$. Thus, the ringing frequency for relatively small values of β is still close to central frequency but somewhat higher. We may also define a measure for the degree of ringing as $\alpha(u_m)\equiv\rho$, the value of the maximum of $\alpha(u)$. ρ is a function of β and reduces from unity at $\beta=0$ essentially to zero when β reaches ~ 0.65. This then is the broadband limit of the box. A wider box than this is essentially broadband.

The quantity ρ appears in a bound that we can derive on the difference in the values of the filtered output separated by a time lag equal to the ringing period $\tau_{\rm ring}$. Let $\Delta c = |c(t+\tau_{\rm ring}) - c(t)|$ and $\Delta\mathbf{q}$ be the difference between the corresponding vectors. Then we have, from Eq. (2.9a) and the Schwarz inequality,

$$\Delta c = |\Delta\mathbf{q}\cdot\mathbf{n}| \le |\Delta\mathbf{q}||\mathbf{n}| \ . \tag{3.17}$$

Since for the box filter the $\mathbf{q}_m$ are of equal length we have to solve an isosceles triangle where the unequal angle is the angle between the concerned vectors. Thus, $|\Delta\mathbf{q}| = 2\sqrt{2k/N}\sin[\theta(\tau_{\rm ring})/2]$ and from Eq. (3.17) and the definition of ρ we obtain

$$\Delta c \le \sqrt{2k/N}\,|\mathbf{n}|(1-\rho)^{1/2} \ . \tag{3.18}$$

If ρ is close to unity as in case (i) the correlation returns more or less to its original value and there is pronounced ringing.

D. The chirp filter

In this subsection we basically derive an equivalent box-filter for the chirp and obtain both ρ and the ringing frequency. To achieve this we compute the angle between the filter vectors and investigate the properties of $\alpha(u)$. The first maximum of $\alpha(u)$ as u increases from zero will provide us with the necessary results. We start with Eq. (2.9b) which reduces to the following due to the seismic cutoff:

$$\mathbf{q}_m = \sum_{n_s}^{N/2-1} \exp[-2\pi imn/N]\tilde{Q}_n \mathbf{f}_n + \text{c.c.} \tag{3.19}$$

The scalar product between the filters is

$$\begin{aligned}\mathbf{q}_m\cdot\mathbf{q}_p &= \frac{2}{N}\sum_{n_s}^{N/2-1}|\tilde{Q}_n|^2\cos[2\pi(p-m)n/N] \\ &\simeq \frac{2}{N}\frac{1}{\Delta^2}\sum_{n_s}^{N/2-1}|\tilde{q}_n|^2\cos[2\pi(p-m)n/N] \\ &\simeq \frac{4}{3}\frac{f_s^{4/3}}{\Delta}\int_{f_s}^{\infty}\frac{\cos(2\pi f\tau)}{f^{7/3}}df \ ,\end{aligned} \tag{3.20}$$

where we have used the relation $\tilde{q}(f_n) \simeq \Delta\tilde{Q}_n$. The integral in Eq. (3.20) is just the autocorrelation function for the correlation as mentioned by Schutz [10]. The norm of the vectors is obtained by setting $\tau = 0$ in Eq. (3.20) and performing the integration. Thus,

$$|\mathbf{q}_m|^2 = \frac{1}{\Delta} \ . \tag{3.21}$$

The angle $\theta(u)$ between the $\mathbf{q}_m$ is then given by

$$\alpha(u) \equiv \cos\theta(u) = \frac{4}{3}\int_1^{\infty}\frac{\cos(2\pi xu)}{x^{7/3}}dx \ , \tag{3.22}$$

where $x = f/f_s$ and $u = f_s\tau$. Note that here we have defined the dimensionless time in units of f_s instead of the central frequency which is not obvious to define for the chirp. The ringing period is that when the $\mathbf{q}_m$ vector returns to the neighborhood of its initial direction and makes the least angle with the initial direction among adjacent vectors. This is when $\alpha(u)$ reaches its second maximum. (The first maximum occurs at $u = 0$). A graph of $\alpha(u)$ versus u is shown in Fig. 3. The second maximum of $\alpha(u)$ occurs at $u = u_m \simeq 0.79$ and the corresponding ringing frequency is

$$f_{\rm ring} \simeq 1.27 f_s \ . \tag{3.23}$$

These parameters agree with an effective box extending from $f_1 = f_s$ to $f_2 \simeq 1.55 f_s$. The box has a bandwidth of about $0.55 f_s$ and central frequency $\sim 1.27 f_s$. The parameter β defined in the last subsection has a value of ~ 0.2. From the discussion in the last section this implies that $f_{\rm ring} \sim f_c$. The intensity of ringing is then

$$\rho = \alpha(u_m) \simeq 0.234 \ . \tag{3.24}$$

Our simulations show that the ringing does not depend on the other parameters of the filter ξ or Φ. This is because the effective box is essentially determined by the frequency dependence of the power spectrum of the filter which falls off as $f^{-7/3}$ irrespective of the filter parameters. Since the $f^{-7/3}$ falloff is calculated in the stationary phase approximation, the above results are valid within this approximation.

The foregoing analysis shows that the ringing is regular noise. It occurs because the chirp filter effectively narrow bands the data which leads to a rotational motion of the filter vectors in $\mathcal{R}^N$. The ringing behavior of the filtered output can be explained by considering the filtered output as a scalar product between a fixed noise vector $\mathbf{n}$ and the filter vectors $\mathbf{q}_m$. The troughs and crests occur according as this scalar product is negative or positive. The ampli-

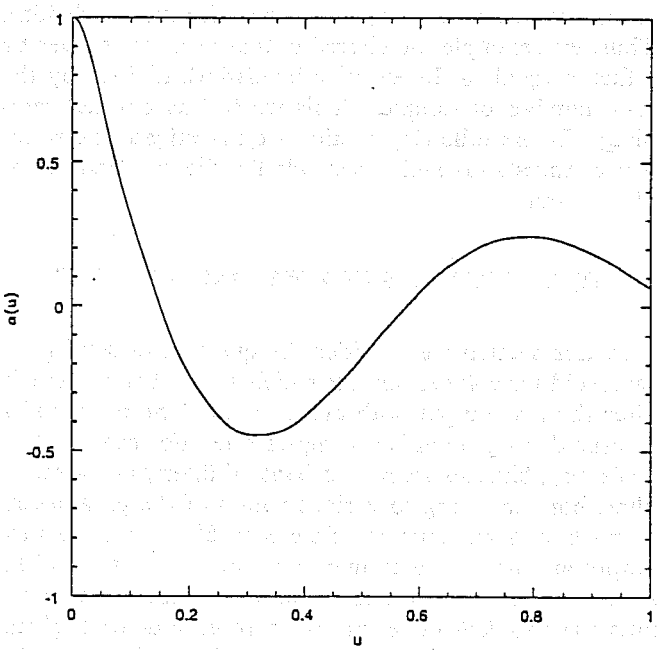

FIG. 3. The figure shows the plot of the normalized autocorrelation $\alpha(u)$ for the chirp filter as a function of the dimensionless time u (in units of f_s^{-1}). As u increases from 0 the second maximum occurs at $u_m \simeq 0.79$ which determines the ringing period. The height of this maximum is $\rho = 0.234$ which when compared with unity gives the intensity of ringing.

tude of the ringing is large when the noise vector lies more or less in the plane of rotation of the $\mathbf{q}_m$ and it is small when the noise vector is roughly orthogonal to the plane. The plane also rotates; slowly, if the filter is narrow band—in one rotation of the $\mathbf{q}_m$ the plane does not change much; or quickly if it is in the broadband regime. When the time lag between two points of the filtered output is large compared with the ringing period, the amplitude for the ringing can differ enormously since the planes can have vastly different orientations at the two time lags.

In the broadband case the motion of the filter vectors can be easily appreciated in the time domain. We have

$$q_{(m+p)l} = q_{(m)l-p} \ , \tag{3.25}$$

where the l component of the filter vector with time lag $m+p$ is the $l-p$ component of the filter vector with time lag m. The time domain components are just "shifted" into a different dimension. The important point is that *there is no rotation* of the filter vectors, hence no *ringing*. There are no "troughs" or "crests" in the usual sense. A signal arriving at some time lag $m\Delta$ is pulled down or pushed up according as the scalar product $\mathbf{q}_m \cdot \mathbf{n}$ is negative or positive.

The important implication of ringing is that there is redundant information in the filtered output. The effective bandwidth of the filter decides the number of independent samples. For the box, it is exactly the number of positive and negative frequencies constituting the box. In case of the chirp the effective box basically portrays this number. This suggests the following strategy for economizing on data storage space and data processing. We can resample the filtered output at a coarser rate, since the filtered output contains the necessary information in a subset of the filtered output. We take this subset to contain as many points as the effective bandwidth. Thus, we resample the filtered output at a rate slower by a factor equal to the effective bandwidth divided by the total number of samples. This we call as optimal sampling. In the following section we investigate the problem of thresholds and thresholds for the optimally sampled output.

IV. THRESHOLDS AND OPTIMAL SAMPLING

In this section we consider the question of setting up threshold for colored outputs which arise when we match filter the raw output with chirp filters. The problem of thresholds in general is a complex one. For example for coalescing binaries we have a bank of filters, i.e., a set of filters corresponding to various values of the parameters ξ and Φ and we must correlate each filter with the raw output and have several filtered outputs to contend with. As there could be covariances between these filtered outputs the problem of setting up a threshold for a given false alarm probability becomes involved. Here we do not consider this extra complication but consider just one filter with fixed values for its parameters ξ and Φ. The only parameter we allow to vary is the time lag τ.

Consider the autocorrelation of the filtered output $c(t)$ when the raw data are only noise. Following Schutz [10], we obtain

$$a(\tau) \equiv \langle c(t)c(t+\tau)\rangle = \int_{f_s}^{\infty} \frac{\cos(2\pi f\tau)}{f^{7/3}} df \ . \tag{4.1}$$

Now $a(\tau)$ resembles a sinc function, which means that the samples for small τ are not statistically independent but become increasingly so for large τ. Schutz suggested defining a "decorrelation time," above which the samples in the filtered output would be uncorrelated. Also, since there are covariances for small time lags there is redundant information in the filtered output and the necessary information should be available in a subset of the filtered output. As there are fewer degrees of freedom in the filtered output, it should be possible to lower the threshold. In this section we examine all these questions and attempt answering them. We perform a detailed analysis and obtain quantitative results on the decorrelation time and optimal sampling rates.

For setting thresholds we first analyze in Sec. IV A the standard case when all the samples of the filtered output are algebraically independent random variables but could be statistically dependent. More specifically, the filtered output C_m, $m=0,1,\ldots,N-1$, are algebraically independent variables; i.e., there does not exist any function or functions connecting different C_m. However, the covariance between different C_m need not vanish. First we treat the simpler case of the uncorrelated output and then treat the correlated case. In Sec. IV B we treat the case when the filtered output samples are linearly dependent because of narrow banding. As mentioned in the last section the number of degrees of freedom is the effective bandwidth of the filtered output which is less than N. Therefore, we write a joint probability density function on the subspace spanned by the Fourier components constituting the narrow band. In Sec. IV we apply this analysis to the box filter and in Sec. IV D we define an effective box for the chirp filter and obtain thresholds.

Let us consider a data train of total length $T_{\rm tot}$. The data train is divided into M segments of length T. As earlier we denote the sampling interval by Δ and the number of points in each segment as N. The total number of points in the entire data train is denoted by $N_{\rm tot}$. We have the following relations between the various quantities:

$$T_{\rm tot} = MT = N_{\rm tot}\Delta \ , \quad N_{\rm tot} = MN \ . \tag{4.2}$$

Typically, $T_{\rm tot} \sim 1$ yr $\sim 3\times 10^7$ sec, $T \sim 30$ sec, $M \sim 10^6$, $N \sim 3\times 10^4$, $N_{\rm tot} \sim 3\times 10^{10}$ and $1/\Delta \sim 1$ kHz. Suppose we can afford a false alarm rate of one per year. Then the false alarm probability P_F is $N_{\rm tot}^{-1}$. For Gaussian noise, each C_m is Gaussian distributed since it is a linear function of Gaussian random variables. By normalizing the filter appropriately we can make the variance of each C_m unity. Assuming zero mean for the noise, each C_m has zero mean. Thus, each C_m is a standard normal variate.

Consider a data segment of N points and let η be the threshold on the filtered output C_m. Let $p_0(\eta)$ be the probability that no C_m crosses the threshold on either

side, i.e., $|C_m|<\eta$, for all m. That is $p_0(\eta)$ is the probability of *no false alarm* for a data segment. Then $P_F=1-p_0(\eta)$ is the probability of at least one false alarm occurring in a data segment. Then to expect one false alarm in the period of time $T_{\rm tot}$, the threshold η must satisfy the equation

$$1-p_0(\eta)=\frac{1}{M}\ . \tag{4.3}$$

The solution to Eq. (4.3) is the required threshold η.

A. Thresholds for the linearly independent samples

Case (a). *Uncorrelated output.* Let us start by considering the trivial case, in order to fix notations. We define

$$\begin{aligned} I(\eta)&\equiv\frac{1}{\sqrt{2\pi}}\int_{-\eta}^{\eta}\exp[-\tfrac{1}{2}x^2]dx \\ &=\mathrm{erf}\left[\frac{\eta}{\sqrt{2}}\right] \\ &\simeq 1-\left[\frac{2}{\pi}\right]^{1/2}\frac{\exp(-\frac{1}{2}\eta^2)}{\eta}\quad\text{for }\eta\gg 1\ . \end{aligned} \tag{4.4}$$

When we assume that all samples C_m are uncorrelated, then the probability that none of the C_m cross the threshold η is

$$p_0(\eta)=I^N(\eta)\ , \tag{4.5a}$$

and the false alarm probability is given by

$$\begin{aligned} P_F(\eta)&=1-I^N(\eta) \\ &\sim N\left[\frac{2}{\pi}\right]^{1/2}\frac{\exp(-\frac{1}{2}\eta^2)}{\eta}\quad\text{for }\eta\gg 1\ . \end{aligned} \tag{4.5b}$$

If we assume a false alarm rate of one in the period $T_{\rm tot}$ then from Eq. (4.3) we have to solve the equation

$$\left[\frac{2}{\pi}\right]^{1/2}\frac{\exp(-\frac{1}{2}\eta^2)}{\eta}\sim(MN)^{-1}=N_{\rm tot}^{-1}\ .$$

For $N_{\rm tot}\sim 3\times 10^{10}$, we get the threshold $\eta\sim 6.6$. We could have got exactly the same result by taking just a single data train of $N_{\rm tot}$ points and assuming all samples to be independent.

Case (b). *Correlated output.* The above analysis can be easily extended at least in principle to accommodate the case of a filtered output which is correlated but algebraically independent. One simply writes a joint probability density function for the N variables C_m:

$$p(\mathbf{C})=\left[\frac{\det\alpha}{(2\pi)^N}\right]^{1/2}\exp[-\tfrac{1}{2}\alpha_{ij}C_iC_j]\ , \tag{4.6}$$

where, C is the N-dimensional vector with components C_m and α_{ij} is the inverse of the matrix $\langle C_iC_j\rangle$. The probability of no false alarm is

$$p_0(\eta)=\int_{\mathcal{H}}p(\mathbf{C})d^NC\ , \tag{4.7}$$

where $\mathcal{H}$ is the N-dimensional volume inside the hypercube defined by

$$\mathcal{H}=\{\mathbf{C}/|C_m|<\eta\ ,\ \ m=0,1,\ldots,N-1\}\ . \tag{4.8}$$

The threshold η is then obtained by solving the equation $1-p_0(\eta)\sim 1/M$.

How is this threshold related to the one of uncorrelated samples in case (a)? If $\alpha_{ij}=\delta_{ij}$ we recover case (a). Let us call the value of this threshold as $\eta_{\rm max}$. The other extreme is that all the samples are perfectly correlated then they are like just one sample and we have to solve the equation $1-\mathrm{erf}(\eta/\sqrt{2})=1/M$. This will yield a lower value of η say $\eta_{\rm min}$. Thus, in the intermediate case the threshold would lie somewhere between $\eta_{\rm min}$ and $\eta_{\rm max}$. For the numbers taken above $\eta_{\rm min}\sim 4.5$ while $\eta_{\rm max}\sim 6.6$.

B. Thresholds for narrow-banded filtered output

In this subsection we consider a narrow-band filter defined as follows: For $n>0$, define $\tilde{Q}_n=0$ for n outside a band $n_0\le n\le k-1$, and we ensure reality by taking $\tilde{Q}_{-n}=\tilde{Q}_n^*$. For this filter we set up the problem of the threshold in general. Since we are defining the threshold based on false alarm probabilities we assume that the raw output consists of noise only. Then the filtered output is given by

$$C_m=\frac{1}{N}\sum_{n_0}^{n_0+k-1}\tilde{N}_n\tilde{Q}_n^*\exp[2\pi imn/N]+\text{c.c.} \tag{4.9}$$

The random variables here are $\tilde{N}_n$ which are complex. But since we are dealing with real quantities such as the filtered output, filter, etc., we split the complex random variables into real and imaginary parts which are now real random variables. The $\tilde{Q}_n$ merely modulate these random variables. We rewrite Eq. (4.9) in terms of these variables defined below:

$$\begin{aligned} &\tilde{N}_{n_0+l}=u_l+iv_l\ , \\ &\tilde{Q}_{n_0+l}=\rho_l\exp(i\psi_l)\ , \\ &\phi_{ml}=\psi_l-2\pi m(n_0+l)/N\ , \\ &C_m=\frac{2}{N}\sum_{l=0}^{k-1}\rho_l[u_l\cos\phi_{ml}+v_l\sin\phi_{ml}]\ . \end{aligned} \tag{4.10}$$

We assume the raw output to be white noise with mean zero and variance unity. Then results from Sec. II imply $\langle\tilde{N}_n\rangle=0$, $\langle\tilde{N}_n\tilde{N}_m^*\rangle=N\delta_{mn}$. The real and imaginary parts of $\tilde{N}_n$, namely, u_l and v_l are Gaussian distributed with means zero and have the following variances and covariances:

$$\begin{aligned} &\langle u_l\rangle=\langle v_l\rangle=0\ , \\ &\langle u_lv_p\rangle=0\ , \\ &\langle u_lu_p\rangle=\langle v_lv_p\rangle=\frac{N}{2}\delta_{lp}\ . \end{aligned} \tag{4.11}$$

The variance of C_m can be computed from Eqs. (4.10) and (4.11):

$$\sigma_C^2 \equiv \langle C_m^2 \rangle = \frac{2}{N}\sum_{l=0}^{k-1} \rho_l^2 \ . \tag{4.12}$$

The joint probability density function defined over $\mathcal{Q}$ in terms of the $2k$ coordinates u_l and v_l, now written as two k-dimensional vectors $\mathbf{u}$, $\mathbf{v}$ is

$$p(\mathbf{u},\mathbf{v}) = \frac{1}{(\pi N)^k}\exp\left[-\frac{1}{N}\sum_{l=0}^{k-1}(u_l^2+v_l^2)\right] \ . \tag{4.13}$$

We can now set up the problem of the threshold as follows: We look upon C_m as linear relations among u_l and v_l. Geometrically, for a given m, the equation $C_m = \text{const}$ represents a hyperplane in $\mathcal{Q}$. This hyperplane is orthogonal to the filter vector $\mathbf{q_m}$ since $C_m = \mathbf{q_m}\cdot\mathbf{n}$. We set the threshold at η times the standard deviation of C_m, i.e., at $\eta\sigma_C$, then the no false alarm condition demands that $|C_m| < \eta\sigma_C$ for all values of m. The relation $|C_m| < \eta\sigma_C$ geometrically represents the region $V_m(\eta)$ between the pair of hyperplanes defined by $C_m = \pm\eta\sigma_C$. Thus,

$$V_m(\eta) = \{(\mathbf{u},\mathbf{v}) \in \mathcal{Q} / |C_m| < \eta\sigma_C\} \ . \tag{4.14}$$

The condition that all C_m have absolute value less than the threshold value is then represented geometrically as the intersection of the regions $V_m(\eta)$, i.e.,

$$V(\eta) = \cap_{m=0}^{N-1} V_m(\eta) \ . \tag{4.15}$$

The no false alarm probability is then obtained by integrating the probability density function over $V(\eta)$:

$$p_0(\eta) = \int_{V(\eta)} p(\mathbf{u},\mathbf{v}) d^k u d^k v \ . \tag{4.16}$$

The threshold is then obtained by solving $1 - p_0(\eta) = 1/M$.

The analysis described above is simple, in principle, but difficult, in general, in practice. The basic reason is that the region $V(\eta)$ can be very awkward for a physical problem. This is so in our application of the coalescing binary filter. The region $V(\eta)$ is determined by the dynamics of the hyperplanes $C_m = \text{const}$ or equivalently the filter vectors $\mathbf{q_m}$. In the next subsection we first treat the case of box filters which is relatively less complex. But even here we found that we could solve the problem exactly when the box consisted of just two frequencies. (We say a box has k frequencies when it consists of k positive and k negative frequencies which are mirror images of each other about the zero frequency.) In all other cases we had to resort to computer simulations.

C. The box filter

We begin with the box filter since this problem is relatively simpler to deal with than the problem of the chirp. The problem of the box with just two frequencies can be treated analytically and this provides invaluable insight for tackling the more advanced problems. We obtain the false alarm probabilities as a function of the threshold level η for the box filter and then go on to discuss optimal sampling rates for the filtered output. In this case the effective number of independent samples N_{eff} is obvious and is equal to the number of positive and negative frequencies constituting the box. If the box consists of k frequencies then, $N_{\text{eff}} = 2k$ and the sampling interval $\Delta_{\text{eff}} = N\Delta/2k$.

The box filter may be defined via the filter vectors $\mathbf{q_m}$. Then Eq. (3.10) defines a box filter with k frequencies. We divide the discussion into two parts: (a) $k=2$ case, (b) $N/2 \geq k \geq 3$ case.

(a) $k=2$. The problem is to be tackled in two stages: first we must determine $V(\eta)$ and second we must integrate the probability density function over $V(\eta)$ to obtain $p_0(\eta)$ and thus the false alarm probability P_F. The subspace $\mathcal{Q}$ is four dimensional and $V(\eta)$ is a four-dimensional volume. The vectors $\mathbf{u}$, $\mathbf{v}$ have two components each and the probability density function $p(\mathbf{u},\mathbf{v})$ is a four-dimensional Gaussian distribution.

In Eq. (3.10), $n = n_0, n_0+1$ and hence $l=0,1$. Setting $\rho_l = 1$, $\theta_l = 0$, $\psi_l = 0$; $l=0,1$ in Eq. (4.10) for the box we obtain

$$C_m = \frac{2}{N}(u_0\cos\phi_{m0} + v_0\sin\phi_{m0} + u_1\cos\phi_{m1} + v_1\sin\phi_{m1}) \ , \tag{4.17}$$

where $\phi_{m0} = -2\pi m n_0/N$, $\phi_{m1} = \phi_{m0} - 2\pi m/N$.

Normally, the numbers N, n_0 are large (as in the case of the chirp filter). In the example considered in the last section $N = 8192$ and $n_0 = 328$. If we assume similar relations here, i.e., $1 \ll n_0 \ll N$ the variables ϕ_{ml} can be considered to vary continuously and we may use the methods of calculus. We consider the family of hyperplanes $C_m = \pm\eta\sigma_C$ and compute their envelope. This we achieve by eliminating ϕ_{ml} between Eq. (4.17) and

$$\frac{\partial C_m}{\partial \phi_{ml}} = 0 \ , \quad l = 0,1 \ . \tag{4.18}$$

In polar coordinates, defined from $u_l + iv_l = r_l \exp(i\kappa_l)$, $l=0,1$, the envelope is given by

$$\frac{2}{N}(r_0 + r_1) = \eta\sigma_C \ . \tag{4.19}$$

This can be understood in the following way: For a given data segment u_0, v_0, u_1, v_1 are fixed and C_m as a function of m is just a sum of two sinusoids which differ by a small beat frequency $2\pi/N$. As m ranges over $0,1,2,\ldots,N-1$ there is exactly one beat and the maximum value of C_m is just the sum of the modulii r_0, r_1 multiplied by the factor $2/N$.

The volume $V(\eta)$ is

$$V(\eta) = \left\{(r_0,\kappa_0,r_1,\kappa_1)/0 \leq \frac{2}{N}(r_0+r_1) \leq \eta\sigma_C; r_0, r_1 \geq 0; 0 \leq \kappa_0, \kappa_1 \leq 2\pi\right\} \ . \tag{4.20}$$

$p_0(\eta)$ is obtained by using Eq. (4.16). We scale the coordinates, $U_l=\sqrt{2/N}u_l$, $V_l=\sqrt{2/N}v_l$, and $R_l=\sqrt{U_l^2+V_l^2}$, so that the new variables have unit variance. Integrating over the angles, we obtain

$$p_0(\eta)=\int R_0 R_1 \exp[-\tfrac{1}{2}(R_0^2+R_1^2)]dR_0 dR_1 , \qquad (4.21)$$

where the integration region is $0\le R_0+R_1\le\eta\sqrt{2}$, $R_0,R_1\ge 0$. This integral yields

$$P_F(\eta)=1-p_0(\eta)$$
$$=\exp(-\eta^2)+\left[\frac{\pi}{2}\right]^{1/2}\eta\exp[-\tfrac{1}{2}\eta^2]\operatorname{erf}\left[\frac{\eta}{\sqrt{2}}\right] . \qquad (4.22)$$

The first term is significant for small η but for large η the second term takes over. The second term resembles the Rayleigh distribution. Figure 4 shows the plot of $-\log_{10}P_F$ versus η. The topmost curve (dashed) shows the $k=2$ case. The lowest (also dashed) curve is for the broadband $k=N/2$: $P_F(\eta)=1-I^N(\eta)$. We have taken here $N=256$ which is small compared with usual case. This allowed us to perform a large number of simulations in a given amount of computer time. The simulated results almost exactly agree with the theoretically obtained curve. The figure shows that the false alarm probability for the $k=2$ case is smaller than in the broadband case at a given value of η; or equivalently the threshold level reduces in the narrowband case. This makes sense intuitively, because the filtered output depends on fewer random variables.

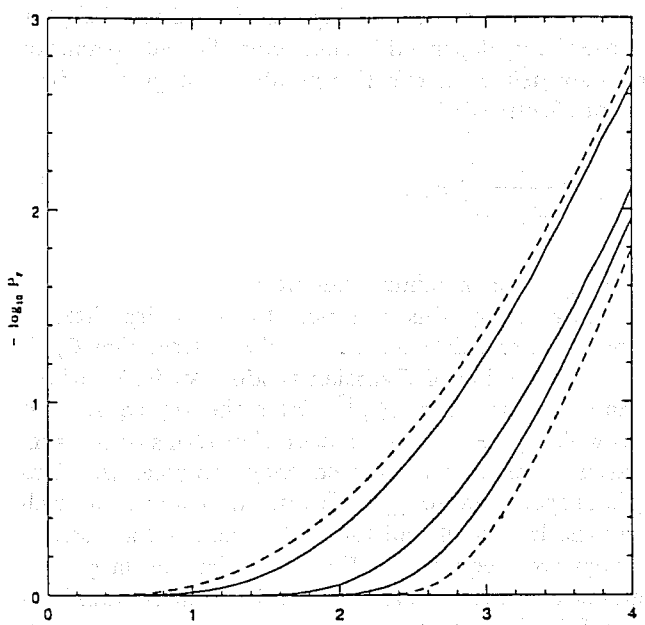

FIG. 4. The figure displays false alarm probability curves where $-\log_{10}P_F$ is plotted vs the threshold η for box filters of various bandwidths. There are five curves corresponding to $k=2$, 3, 10, 32, and 128. $N=256$ for all cases, and the least frequency is $n_s=101$ for all but the broadband case $k=128$. The various curves can be picked out by the property that as k increases they move "down" in the figure implying that the false alarm probability P_F increases for a given value of the threshold. The topmost $k=2$ and the bottommost $k=128$ curves are dashed and have been obtained analytically.

The filtered output has just four degrees of freedom in this case so that much of the information in it is redundant. $N_{\rm eff}=4$ here and we may resample the output at just four equidistant points of the filtered output. We may again construct a false alarm probability curve with the resampled output. We find that this agrees with the false alarm probability $P_F(\eta)=1-I^4(\eta)$. This shows that the four samples are independent random variables.

(b) $k\ge 3$. The situation becomes more difficult for more than two frequencies. The basic reason for this is that there is more than one beat frequency but a single parameter m. There does not exist in general a value of m for which the modulii of the complex noise components add up, i.e., the maximum value of C_m is not necessarily the sum of the modulli, but is in general less. Alternatively, we cannot eliminate the angular dependence from the envelope and hence from the volume $V(\eta)$. This makes the volume $V(\eta)$ awkward and the integration becomes intractable. We had therefore to resort to simulations and Fig. 4 shows the results. The solid line curves are the false alarm probability curves for the values $k=3,10,32$. The curves move "downwards" as the value of k increases which means the false alarm probability increases as k increases, for a given η. Also the curves seem to converge rapidly towards the broadband. The $k=3$ curve is well separated from the $k=2$ curve and the $k=32$ curve is very close to the broadband curve ($k=128$). Further our simulations show that the false alarm probability curves are insensitive to the value of n_0, i.e., the position of the box.

We also investigated the resampling of the filtered output taking $2k$ equidistant samples. We computed the false alarm probability curves with the resampled filtered output and find that they agree very well with the curves given by $P_F(\eta)=1-[\operatorname{erf}(\eta/\sqrt{2})]^{2k}$, which shows that the $2k$ samples are independent random variables.

D. The chirp filter

We divide the discussion here into two parts: (a) false alarm probabilities, (b) optimal sampling.

(a) *False alarm probability.* The problem for the chirp is even harder to tackle analytically. We do not attempt it but try to solve the problem by performing computer simulations. We choose the following parameters for the simulations: $N=8192$, $\Delta=4\times10^{-4}$ sec, thus $T\simeq 3.28$ sec. The seismic cutoff is at $f_s=100$ Hz and hence $n_s=[f_sT]=327$. The length of the data segment constraints the duration of the chirp signal and hence the chirp mass. We consider $\xi\le 3$ sec, which means that the chirp mass must be greater than one solar mass. One of the aspects of our analysis is to normalize the variance of C_m to unity, which is done by dividing out by the factor σ_C at some stage in the computation of the correlation. From Eq. (4.12) σ_C is the norm of q_m which from Eq. (3.21) is $\Delta^{-1/2}=50$ for the parameters chosen.

The simulations are performed on the lines of the box

filter. We generate white noise with unit variance. This is the raw output. We generate the chirp filter with the chirp wave form from $t=0$ to $t=\xi$ sec with its instantaneous frequency at $t=0$ equal to 100 Hz. For $\xi \le t \le T$ the filter is padded with zeros. We then correlate the raw output with the chirp filter to obtain the filtered output. We have used routines from Ref. [13] for our simulations. We took filters with several values of $\xi=0.5$, 2.0, 3.0 sec and we find that the results are insensitive to the value of ξ. The false alarm probability is slightly less than the broadband case of all 8192 uncorrelated samples. The threshold is reduced by about 0.05 of the standard deviation of the filtered output. Figure 5(a) shows the two curves: the solid curve is for the chirp filter while the dashed curve is for the broadband false alarm probability $P_F(\eta)=1-[\mathrm{erf}(\eta/\sqrt{2})]^{8192}$.

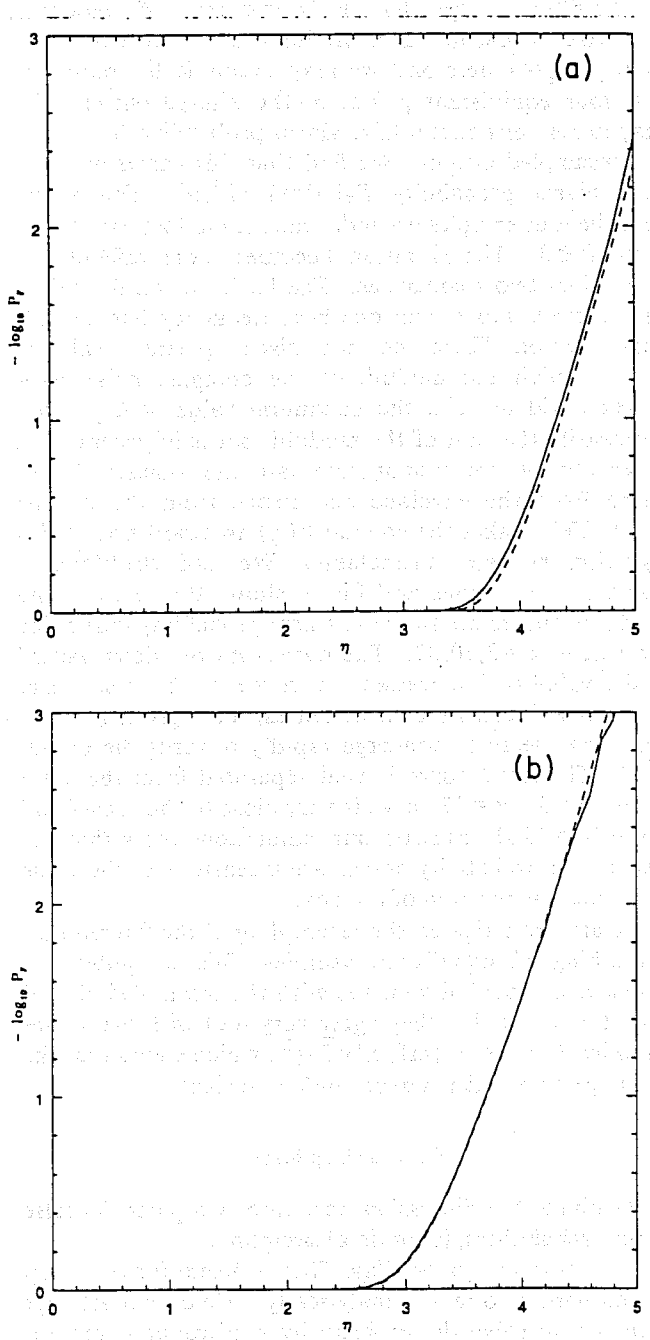

FIG. 5. (a) The figure shows the false alarm probability curve (solid line) $-\log_{10}P_F$ vs the threshold η for the chirp with $\xi=3$ sec, $N=8192$, $\Delta^{-1}=2.5$ kHz, and $f_s=100$ Hz. The dashed curve is a reference curve for the false alarm probability if all samples are assumed to be independent. The threshold is seen to be slightly lowered by about 0.05. (b) The figure shows the false alarm probability curve (solid line) with the same parameters as in 4 (a) but for the resampled output. The resampling rate is 150 Hz, 1.5 times the seismic cutoff frequency. Since $N=8192$ the optimal number of samples is $N_{\rm eff}=491$. We observe that this curve coincides with the reference (dashed) curve drawn for $N_{\rm eff}$ independent samples.

(b) *Optimal sampling.* Since there is difference in the false alarm probabilities in Fig. 5(a) it is clear that we are oversampling the filtered output. The question is by what factor? Can we propose an effective box size for the chirp filter so that the false alarm probability curve for the slower sampled output coincides with the broadband curve for the resampled output? In this subsection we attempt to answer these questions.

We observe that C_m is a sum of independent Gaussian variables and the quantity that affects the false alarm probability curves is basically the variance of C_m. Let us consider the following situation. Let

$$C=\sum_{i=1}^{n} X_i \ , \tag{4.23}$$

where X_i are independent Gaussian random variables with mean zero and variances σ_i^2, respectively. Then the variance of C, namely, $\sigma_C^2=\sum_{i=1}^{n}\sigma_i^2$. If each $\sigma_i=1$ then $\sigma_C^2=n$ the number of random variables. If only k of the X_i had unit variance and others had vanishing variance then $\sigma_C^2=k$, the number of degrees of freedom of C. But suppose the σ_i depended in some complicated manner on i, we then define the effective number of degrees of freedom, or effective dimension by

$$N_{\rm eff}=\frac{1}{\sigma_{\rm max}^2}\sum_{i=1}^{n}\sigma_i^2 \ , \tag{4.24}$$

where $\sigma_{\rm max}$ is the maximum among σ_i.

We now apply this criterion to the chirp filter T seconds long and determine $N_{\rm eff}$. We observe that C_m is linear combination of Gaussian random variables whose variances are scaled as $|\tilde{Q}_n|^2$. Here the key equation is (4.12) which gives the variance of C in terms of the sum of variances of the various frequency components. Thus $N_{\rm eff}$ is proportional to $\sum\rho_l^2$. We must, however, normalize by the largest ρ_l^2 and take into account the positive and negative frequencies. For the chirp the largest ρ_l occurs at the seismic cutoff, where the Fourier transform of the filter has maximum amplitude. Thus, the maximum of ρ_l is ρ_0 and, therefore,

$$N_{\rm eff}=2\sum_{l=0}^{N/2-n_s-1}\left[\frac{\rho_l}{\rho_0}\right]^2 . \tag{4.25}$$

The $\rho_n \sim n^{-7/6}$ where $n=n_s+l$ and we obtain

$$N_{\rm eff}=2\sum_{n_s}^{N/2-1}\left[\frac{n_s}{n}\right]^{7/3}$$
$$\simeq 2T\int_{f_s}^{\infty}\left[\frac{f_s}{f}\right]^{7/3}df$$
$$=1.5f_sT\,. \tag{4.26}$$

Thus the optimal sampling rate is $\Delta_{\rm eff}^{-1}=N_{\rm eff}/T$ which is $1.5f_s$. The effective bandwidth of the positive frequency box is $N_{\rm eff}\Delta^{-1}=0.75f_s$. For $f_s=100$ Hz, the optimal sampling rate is 150 Hz and the number of samples in the optimally sampled output is $N_{\rm eff}\sim 491$. Figure 5(b) shows that the false alarm probability curves merge together for the resampled output and broadband curve defined by $P_F(\eta)=1-[\mathrm{erf}(\eta/\sqrt{2})]^{N_{\rm eff}}$.

V. CONCLUSION

We have shown that the ringing of the correlation is due to the effective narrow banding by the matched filter. In the geometrical picture, the filter vectors of a narrow-band filter execute a rotational motion as the filter is steadily translated in time. Since the correlation in this formalism is just the scalar product of the filter vectors with the noise vector, which is a fixed vector for a given data segment, the correlation rings. This also shows that the ringing is regular noise and the problem of the signal peak getting pulled down in the trough is not different, in principle, from the broadband one where the signal peak could also be pushed up or down depending on the noise at that particular time lag. The analysis shows that the ringing frequency is roughly the same as the central frequency of the box filter if the box is not too wide, i.e., if $\beta\leq 0.4$. In case of the chirp the expected value of the ringing frequency is $1.27f_s$. The degree of ringing is measured by the parameter ρ which lies between zero and unity; $\rho=0$ implies no ringing, while $\rho=1$ implies strong ringing. We use ρ to obtain a bound on the difference in the values of the correlation separated by a wavelength of the ring. For the chirp filter we show that $\rho=0.234$.

We then investigate the problem of thresholds for detection when the filtered output is narrow banded. We set up a formalism for this purpose and show that the false alarm probability can be obtained by performing an integration over a volume in the subspace spanned by the frequency components of the narrow band. However, these volumes turn out to be awkward in general, except in the case of a box filter having two frequencies. For box filters with more than two frequencies and for the chirp filter we resort to simulations to obtain the thresholds as a function of the false alarm probabilities. The false alarm probability increases as the width of the box increases. We also find that these curves are insensitive to the position of the box in the frequency space, i.e., if the bandwidth is kept fixed but its central frequency is varied. For the chirp we find that the false alarm probability curve does not depend on its parameters and lies quite close about 0.05 standard deviations away from the curve obtained by assuming all samples to be independent, i.e., the threshold is lowered by this amount for the colored output.

Lastly we consider optimal sampling rates and find that for detection purposes the number of independent samples $N_{\rm eff}$ in the filtered output is essentially the bandwidth of the effective box. For the box filter it is just the number of real frequencies, both positive and negative, that constitute the box. For the chirp filter we define an effective box based on the analysis of the variance of the correlation and find that this leads to a sampling rate of $1.5f_s$. For the example considered ($f_s=100$ Hz) the sampling rate is 150 Hz. From simulations it is found that the false alarm probability curves for the resampled output agree with the false alarm probability curve obtained by assuming the resampled points to be statistically independent.

Note that, for the box filter, another way to approach the sampling problem is to *heterodyne*. By moving the box to zero frequency, the usual Nyquist theorem tells us what the sampling rate we must use. The Nyquist rate is exactly the rate we get for the box. Therefore, for the colored chirp data, we are effectively establishing a Nyquist sampling rate.

Here we have treated the simpler problem where a single chirp filter is taken and the covariances are considered for various time lags. But the actual problem of detection involves a host of filters and there would exist covariances between filtered outputs of different filters. This problem we wish to investigate in the future.

ACKNOWLEDGMENTS

The authors would like to thank Dr. David Nicholson for his invaluable help in computing and simulations. One of the authors (S.V.D.) wishes to thank the SERC for financial support.

[1] J. Weber, Phys. Rev. **117**, 306 (1960).
[2] K. S. Thorne, in *300 Years of Gravitation*, edited by S. W. Hawking and W. Israel (Cambridge University Press, Cambridge, England, 1987).
[3] A. Abramovici, W. E. Althouse, R. W. P. Drever, Y. Gursel, S. Kawamura, F. J. Raab, D. Shoemaker, L. Sievers, R. E. Spero, K. S. Thorne, R. E. Vogt, R. Weiss, S. E. Whitcomb, and M. E. Zucker, Science **256**, 325 (1992).
[4] R. E. Vogt, *Sixth Marcel Grossmann Meeting*, Proceedings, Kyoto, Japan, 1991, edited by H. Sato and T. Nakamura (World Scientific, Singapore, 1992).
[5] C. Bradaschia, *et al.* Nucl. Instrum. Methods Res. Sect. A **289**, 518 (1990).
[6] J. Hough, B. J. Meers, G. P. Newton, N. A. Robertson, H. Ward, G. Leuchs, T. M. Niebauer, A. Rudiger, R. Schilling, L. Schnupp, H. Walther, W. Winkler, B. F. Schutz, J. Ehlers, P. Kafka, G. Schafer, W. M. Hamilton, I. Schutz, H. Welling, J. R. J. Bennett, I. F. Corbett, B. W. H. Edwards, R. J. S. Greenhalgh, and V. Kose, "Proposal for a Joint German-British Interferometric Gravitational Wave

Detector," Max Planck Institute for Quantum Optics, MPQ Report No. 147, GWD/137/JH(89), 1989 (unpublished).

[7] R. Narayan, T. Piran, and A. Shemi, Astrophy. J. **379**, L17 (1991); E. S. Phinney, *ibid.* **380**, L17 (1992).

[8] There are many books that discuss the theory of hypothesis testing, the method of maximum likelihood and matched filtering. We quote below a few standard text books. C. W. Helstrom, *Statistical Theory of Signal Detection*, 2nd ed. (Pergamon, London, 1968); K. S. Shanmugan and A. M. Breiphol, *Random Signals: Detection, Estimation and Data Analysis* (Wiley, New York, 1989).

[9] W. J. Watkins, Ph.D. thesis, University of Wales, College Cardiff, 1991.

[10] B. F. Schutz, in *The Detection of Gravitational Radiation*, edited by D. Blair (Cambridge University Press, Cambridge, 1991).

[11] B. S. Sathyaprakash and S. V. Dhurandhar, Phys. Rev. D **44**, 3819 (1991).

[12] L. S. Finn, Phys. Rev. D **46**, 5236 (1992).

[13] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes, The Art of Scientific Computing* (Cambridge University Press, Cambridge, England, 1986).

[14] C. Bradaschia, E. Calloni, M. Cobal, R. Del Fabbro, A. Di Virgilio, A. Giazotto, L. E. Holloway, H. Kautzky, B. Michelozzi, V. Montelatici, D. Passuello, and W. Velloso, in *Gravitation: A Banff Summer Institute*, edited by R. Mann and P. Wesson (World Scientific, Singapore, 1991).

[15] B. F. Schutz (unpublished).

[16] A. Rudiger and R. Schilling (private communication).

[17] P. C. Peter and J. Mathews, Phys. Rev. **131**, 435 (1963).

[18] B. F. Schutz and M. Tinto, Mon. Not. R. Astron. Soc. **224**, 131 (1987).

[19] S. V. Dhurandhar and M. Tinto, Mon. Not. R. Astron. Soc. **234**, 663 (1988).

[20] Y. Gursel and M. Tinto, Phys. Rev. D **40**, 3884 (1989).

[21] L. S. Finn and D. F. Chernoff, Phys. Rev. D **47**, 2198 (1993).