MATHEMATICS OF GRAVITATION
PART II, GRAVITATIONAL WAVE DETECTION
BANACH CENTER PUBLICATIONS, VOLUME 41
INSTITUTE OF MATHEMATICS
POLISH ACADEMY OF SCIENCES
WARSZAWA 1997

GRAVITATIONAL WAVES FROM COALESCING BINARIES: A HIERARCHICAL SIGNAL DETECTION STRATEGY

S. D. MOHANTY and S. V. DHURANDHAR

Inter-University Centre for Astronomy and Astrophysics
Post bag 4, Ganeshkhind, Pune-7, India
E-mail: mohanty@iucaa.ernet.in, sdh@iucaa.ernet.in

Abstract. The detection of gravitational waves from coalescing compact binaries would be a computationally intensive process if a single bank of template waveforms (i.e., a one step search) is used. We present, in this paper, an alternative method which is a hierarchical search strategy involving two template banks. We show that the computational power required by such a two step search, for an on-line detection of the one parameter family of Newtonian signals, is 1/8 of that required when an on-line one step search is used. This reduction is achieved when signals having a strength of ~ 8.8 times the noise r.m.s. value are required to be detected with a probability of ~ 0.95 while allowing for not more than one false event per year on the average. We present approximate formulae for the detection probability of a signal and the false alarm probability. Our numerical results are specific to the noise power spectral density expected for the initial LIGO.

1. Introduction. The inspiral, due to Gravitational radiation reaction, of a binary composed of compact massive objects (Neutron stars or Black Holes) will produce a gravitational wave signal [3] which, during the last few minutes before merger, will lie within the bandwidths of upcoming laser interferometric detectors like the LIGO [1], VIRGO [2], and GEO600. The waveform of this signal can be computed with enough accuracy to allow pattern matching techniques, like matched filtering, to considerably enhance the signal to noise ratio [3, 4]. Therefore, it should be possible to detect such events upto a large distance and hence observe a significant event rate.

For the detection of signals from coalescing compact binaries, it will be required to correlate the detector output with a bank of template waveforms since the signals will not have a unique waveform but will depend on a number of parameters characterizing the source, like the masses of the components, their spins etc. This discrete set of templates

---

1991 *Mathematics Subject Classification*: 83C35, 94A13.

The paper is in final form and no version of it will be published elsewhere.

must be chosen such that even those signals which are not represented exactly in the set are detected with a significant probability (typically, 0.95). At the same time constraints imposed by the available computing power must be met. A formalism which leads to an optimum bank of templates was set up in [6, 7] and is now known as the S-D formalism. And a search strategy using a single bank of templates is called a *one step search*.

It was shown in [8, 9] that the use of a single template bank would be computationally expensive when the post-Newtonian effects are incorporated. An alternative strategy could be to use several template banks in a hierarchy such that information provided by a coarsely spaced bank of templates at a lower level is used to restrict the search region for a more finely spaced template bank at a higer level.

The use of such a strategy, known as a *hierarchical search*, has been proposed a number of times [8, 9, 10]. However, a detailed formalism for the same has not been given so far in the context of the detection of gravitational wave signals. We present in this paper a rigorous formalism to describe a two step hierarchical search and a first estimate of the numbers involved.

The main result of this paper is that the computational power required for an on-line two step search can be upto a factor of eight smaller than that for an on-line one step search. This happens when a detection probability of $\simeq 0.95$ is sought for signals having a signal to noise ratio (s.n.r.) of 8.8σ and the false alarm is kept so that there is, on the average, not more than one false event per year. This factor of eight is, however, not the last word since although our formalism can yield higher factors (~ 13) for slightly lower strengths, the assumption of statistical independence of correlation outputs, used in deriving some of the formulae, breaks down in these cases. We use the family of Newtonian waveforms for our computations and the noise power spectral density expected for the initial LIGO. However, this formalism can be easily extended to post-Newtonian waveforms as well as a larger number of intermediate stages.

The rest of the paper is organized as follows. The main objective in Section 2 is the presentation of formulae for the detection probability of a signal and the false alarm probability. We start with section 2.1 where we briefly describe the noise and signals that will be used in this paper. We obtain the test statistic relevant to this choice of signals and noise in section 2.2. We discuss the distribution functions of this test statistic in section 2.3. An approximate expression for detection probability is obtained. The false alarm probability is also calculated in the context of a one step search which is discussed in section 2.4.

Section 3 is devoted to the two step hierarchical search. In section 3.1, a general formalism, and associated set of notations, is introduced to describe a two step hierarchical search. In section 3.2 It is shown, that there exists an optimum set of spacings and thresholds which minimizes the computational requirement. An algorithm to obtain this optimum solution is presented. In section 3.3 the computing power required for an on-line two step search is estimated and numerical results are presented. We conclude with Section 4.

2. Preliminaries. Some of the results in the following have been obtained from or compared with Monte Carlo simulations. For these simulations, we have mainly used the

Gaussian random number generator, G05FDF, provided in the NAg library of numerical routines. Wherever possible, the results have been checked for consistency with those obtained using the routine GASDEV provided in Numerical recipes [13].

2.1. The noise power spectral density and the Newtonian waveform. It is expected that the use of environmental monitors and coincidence checks between different detectors will allow the removal of non-Gaussian components from the detector noise. The remaining Gaussian component is expected to be stationary over a time scale of hours. The power spectral density, $S_n(f)$, of the Gaussian noise component will have steeply rising parts at both the low and high frequency ends [1, 12]. The sensitive bandwidth for the initial LIGO would be from [12] $f_a = 40$ Hz to $f_c = 1$ kHz and the output would have to be bandlimited to this band.

The lowest order approximation to the waveform from a coalescing compact binary is provided by the quadrupole formalism [11]. The response of an interferometric detector to such a waveform can be written as [7],

$$h(t;\, \mathcal{A}, t_a, \xi, \Phi) = \mathcal{A}\, a(t-t_a,\, \xi)\cos(\phi(t-t_a,\, \xi)+\Phi)\ . \tag{1}$$

This is the family of signals that is to be detected.

The parameter $\mathcal{A}$ takes into account the distance to the binary as well as various geometrical factors [3, 14, 18, 19] and is effectively a constant for the signal durations considered here. The other, time dependent, part of the amplitude is, $a(t,\, \xi) = [\,1 - t/\xi\,]^{-1/4}$. The phase of the waveform can be expressed as $\phi(t,\, \xi) = 2\pi\int_0^t f(t', \xi)dt'$, where the integrand is the instantaneous frequency of the signal, $f(t,\, \xi) = f_a a(t,\, \xi)^{3/2}$.

Thus, the waveform is a *chirp* whose amplitude and instantaneous frequency increase with time. The rate at which the instantaneous frequency increases is governed by the parameter ξ, called the *chirp time*,

$$\xi = 34.54\left[\frac{\mathcal{M}}{M_\odot}\right]^{-5/3}\left[\frac{f_a}{40\mathrm{Hz}}\right]^{-8/3}\ \mathrm{sec}\ , \tag{2}$$

where $\mathcal{M}$, the *chirp mass*, is the following combination of the reduced mass μ and the total mass M of the binary, $\mathcal{M} = (\mu^3 M^2)^{1/5}$. The low frequency cutoff makes the amplitude of the signal negligible when its instantaneous frequency lies below f_a. The time at which $f(t, \xi) = f_a$ is denoted by t_a which can be taken as the *time of arrival* of the signal. The phase of the signal at t_a is denoted by Φ.

The high frequency cutoff, f_c, will also force the amplitude to a negligible value for instantaneous frequencies beyond f_c. Besides this the nature of the waveform will change when the compact bodies plunge towards each other once the last stable orbit is reached. The waveform has, therefore, an effectively finite duration which, to a very good approximation, is ξ itself.

We can also write the waveform in Eq. (1) as

$$h(t;\, \mathcal{A}, t_a, \xi, \Phi) = \mathcal{A}\, h_0(t-t_a;\, \xi)\cos(\Phi) + \mathcal{A}h_{\frac{\pi}{2}}(t-t_a;\, \xi)\sin(\Phi)\ , \tag{3}$$

where, $h_0(t;\, \xi) = a(t,\, \xi)\cos(\phi(t,\, \xi))$ and $h_{\frac{\pi}{2}}(t;\, \xi) = a(t,\, \xi)\cos(\phi(t,\, \xi)+\pi/2)$.

2.2. The test statistic and its computation. The one step search strategy for the detection of the waveform described above will be implemented as follows [7].

The detector output, $x(t)$, will be sampled at the Nyquist frequency of ~ 2 kHz to give the time series ($x_i = x(i\Delta)$; $i = 0, \ldots, N-1$) where Δ is the sampling interval. The time series should of course be longer than the duration of the longest template or equivalently, the largest chirp time, ξ_{max}. For every value of ξ included in the template bank, two correlations will be computed (ξ_m belongs to the template bank): $C_0(i\Delta,\, \xi_m) = \sum_j x_{j+i} q_{0,j}(\xi_m)$ and $C_{\frac{\pi}{2}}(i\Delta,\, \xi_m) = \sum_j x_{j+i} q_{\frac{\pi}{2},j}(\xi_m)$. The time series q_0 and $q_{\frac{\pi}{2}}$ are the templates *matched* [5] to the quadrature components h_0 and $h_{\frac{\pi}{2}}$. Their Discrete Fourier transform [15, 16] components are given by $\hat{q}_{0,j}(\xi) = \mathcal{N}_h \hat{h}_{0,j}(\xi)/S_n(\,j/N\Delta\,)$ and $\hat{q}_{\frac{\pi}{2},j}(\xi) = \mathcal{N}_h \hat{h}_{\frac{\pi}{2},j}(\xi)/S_n(\,j/N\Delta\,)$. $\mathcal{N}_h$ is a normalization constant defined later on. The quadrature components have the following properties,

$$\sum_j q_{0,j}(\xi) h_{0,j}(\xi) \quad \approx \quad \sum_j q_{\frac{\pi}{2},j}(\xi) h_{\frac{\pi}{2},j}(\xi)\ , \tag{4}$$

$$\left[\sum_i q_{0,i}(\xi) h_{\frac{\pi}{2},i}(\xi)\right]^2 \quad \ll \quad \sum_j q_{0,j}(\xi) h_{0,j}(\xi) \sum_k q_{\frac{\pi}{2},k}(\xi) h_{\frac{\pi}{2},k}(\xi)\ , \tag{5}$$

Correlations can be computed very efficiently using the Fast Fourier Transform (FFT) [15]. But in the time series of a template with chirp time ξ_m, samples for $i > \xi_m/\Delta$ will be zero since the template has a finite duration of ξ_m. Therefore when a correlation between the template and the detector output is taken using an FFT, only the first $N - \xi_m/\Delta$ samples will be the result of a linear correlation [4]. It is desirable to have equal lengths of correlations for every template. Hence, only the first $N_p = N - \xi_{max}/\Delta$ samples will be retained in each correlation and the rest discarded. We call N_p, the *padding* for the template bank. A useful figure for N_p is $\sim 5 \times 10^5$ corresponding to $N = 256 \times 2048$ and $\xi_{max} = 32.0$ sec.

The two correlations obtained above will then be squared and added to produce another time series whose i^{th} sample we denote as $X_i(\xi_m)$ ($i = 0, \ldots, N_p$ now), $X_i(\xi_m) = [C_0^2(i\Delta,\, \xi_m) + C_{\frac{\pi}{2}}^2(i\Delta,\, \xi_m)]^{1/2}$. We call this time series the *rectified output* of a template ξ_m. We denote by λ_m the maximum value among the samples $X_i(\xi_m)$. The above process is repeated for all the chirp times belonging to the template bank and the set $\{\lambda_m\}$ is obtained.

Finally, a *test statistic*, Λ, is constructed as,

$$\Lambda = \max_m \{\lambda_m\}\ , \tag{6}$$

and Λ is compared with a threshold η. If $\Lambda > \eta$ then a detection is anounced or else not.

We now define some quantities which will be of use later on. We call the rectified output of a template when $x(t)$ consists of only a signal and no noise as the *processed form* of the signal produced by the template. The maximum value that any processed form of a signal can have is called the *strength*, S, of the signal [7]. This value is attained only for a template with the same chirp time as that of the signal. For a signal with amplitude $\mathcal{A}$, $S = \mathcal{A}/\mathcal{N}_h$, where $\mathcal{N}_h$ is the normalization constant for a template with the *same* chirp time as that of the signal. When the template chirp time is not the same as that of the signal, the maximum of the processed form will be reduced. We call the

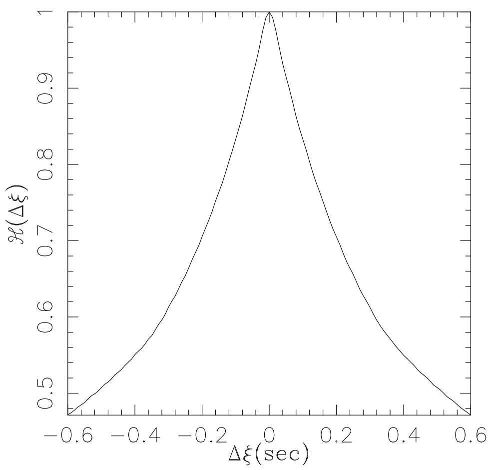

Figure 1: The intrinsic ambiguity function, $\mathcal{H}(\Delta\xi)$, for initial LIGO noise power spectral density.

maximum of the processed form in such a case as the *observed strength*, S_{obs}, of the signal in that template.

Given a signal with $S = 1$ and a chirp time ξ_s, we define the *intrinsic ambiguity* function $\mathcal{H}(\xi_t, \xi_s)$ as the observed strength of the signal in template ξ_t. When $S \neq 1$,the observed strength can be obtained as

$$S_{obs} = S\,\mathcal{H}(\xi_t, \xi_s)\ . \tag{7}$$

It is easy to show, using the stationary phase approximation [6], that $\mathcal{H}(\xi_t, \xi_s)$ depends only on $|\Delta\xi|$ and not on ξ_t and ξ_s separately. This behavior is replicated by the exact intrinsic ambiguity function also. However, for binaries with massive components the plunge cutoff frequency is small. This effectively reduces the signal bandwidth and leads to a wider intrinsic ambiguity. For simplicity, we neglect this effect and take the intrinsic ambiguity to be independent of its location. Our final results will not be affected much by this approximation. A plot of the intrinsic ambiguity function is shown in Fig. 1.

2.3. False alarm and detection probability. The random variables $C_0(i\Delta, \xi_m)$ and $C_{\frac{\pi}{2}}(i\Delta, \xi_m)$ are linear combinations of the time samples of $x(t)$. Since the noise in $x(t)$ was assumed to be Gaussian random process, $C_0(i\Delta, \xi_m)$ and $C_{\frac{\pi}{2}}(i\Delta, \xi_m)$ will, therefore, be Gaussian random variables. The mean values of C_0 and $C_{\frac{\pi}{2}}$ would, in general, be non-zero when a signal is present but would be zero in the absence of a signal. Using the properties of q_0 and $q_{\frac{\pi}{2}}$ stated in the previous section, it is easy to show that the C_0 and $C_{\frac{\pi}{2}}$ are statistically independent. Their variance can be made equal to unity by choosing $\mathcal{N}_h$ suitably and this *defines* $\mathcal{N}_h$.

Thus, it follows that $X_i(\xi_m)$ has a Rayleigh density when a signal is absent and a Rician density when a signal is present [5, 17]. These densities have the following forms: $R(z) = z\exp(-z^2/2)$ for the Rayleigh and $Ri(z, d) = z\exp(-(z^2 + d^2))I_0(dz)$ for the Rician ($I_0(z)$ is the modified bessel function of order zero). The quantity corresponding to d, in the case of $X_i(\xi_m)$, will be given by the i^{th} sample of the processed form of the signal.

If we make the assumption that the set of samples in a rectified output is a statistically independent set, then the cumulative distribution function of λ_m would be a product of the cumulative distribution functions of the samples. We denote the distribution functions of λ by $F_{0,m}(z)$ when a signal is absent and by $F_{1,m}(z;\,\bar{\mu}_s)$ when a signal with parameters $\bar{\mu}_s$ is present.

Of course, the assumption of statistical independence is not strictly correct. However, we found from Monte Carlo simulations that the true $F_{0,m}$ can still be fit by a product of distributions provided the number of samples is reduced from N_p to an effective value $N_{eff} = 0.7N_p$. The ratio N_{eff}/N_p is independent of N_p and ξ_m. The latter property allows us to drop the subscript 'm' and we get,

$$F_{0,m}(z) = F_0(z) = \left(\int_0^z R(x)dx\right)^{N_{eff}} \simeq \exp(-N_{eff}\exp(-\frac{z^2}{2}))\ . \tag{8}$$

Having got $F_0(z)$, we make the assumption that the set $\{\lambda_j\}$ is a statistically independent set. The false alarm probability Q_0 for a given template bank and threshold η can then be expressed as,

$$Q_0 = 1 - (\,F_0(\eta)\,)^{N_T} \simeq 1 - \exp(-N_T N_{eff}\exp(-\frac{\eta^2}{2}))\ , \tag{9}$$

where N_T is the number of templates in the template bank. Monte Carlo simulations show that the threshold, computed for a given false alarm, is actually quite insensitive to the presence of statistical correlations. Thus, the assumption made above is not very restrictive.

In the presence of a signal and for the kind of signal strengths encountered in our analysis, the following approximation is well supported by Monte Carlo simulations. The maximum over all the rectified outputs (i.e., Λ) is localized among those samples where the processed form of the signal is large ($\gtrsim 5.0$). In fact, it is possible to get good fits to Monte Carlo results if only the two templates nearest to a signal (i.e., having chirp times on either side of the signal chirp time) are considered and further, within each of the two rectified outputs, only that sample is considered at which the processed form of the signal is maximum. Under this approximation and the assumption that the two maxima are statistically independent, it can be shown that the detection probability Q_d is given by,

$$Q_d = 1 - \int_0^\eta Ri(x, S_m)dx \int_0^\eta Ri(x, S_{m+1})dx\ , \tag{10}$$

where ξ_m, ξ_{m+1} are the templates between which the signal lies and S_m, S_{m+1} are the respective observed strengths. The assumption of statistical independence becomes more erroneous for smaller values of $\xi_m - \xi_{m+1}$. We restrict ourselves to $\xi_m - \xi_{m+1} \geq 0.030$ sec for which the Q_d calculated from Eq. (10) is off by $\sim 10\%$ from the Monte Carlo results.

2.4. One step search. We will now reformulate the S-D formalism for a one step search in terms of detection probability.

We begin by stating our *one step template placement criteria* : The bank of templates should be chosen in such a way that **(i)** every waveform, having a strength S greater than a given minimum strength S_{min}, should have a detection probability greater than a given minimum detection probability $Q_{d,min}$, and **(ii)** The false alarm should stay below a specified level, $Q_{0,max}$. A solution in terms of η and $\{ \xi_j \}$, satisfying both criteria, need not always exist. For instance, a signal having $S = S_{min} = 6.0$ will not be detected with a detection probability of $Q_{d,min} = 0.95$ if the false alarm is kept such that there is only false event, on the average, in a year.

The detection probability can be expected to be the smallest for signals having a strength S_{min} and chirp time $\xi = (\xi_m + \xi_{m+1})/2$ for $\xi_m \in \{ \xi_j \}$. Such signals will have a detection probability given by $Q_d(\eta;\, S_{min}, (\xi_m + \xi_{m+1})/2)$, which can be calculated using Eq. (10). To satisfy criterion (i) above, all that needs to be done, given a threshold η, is to ensure that all such minimum detection probability signals have, $Q_d(\eta;\, S_{min}, (\xi_m + \xi_{m+1})/2) = Q_{d,min}$. It follows from the location independence of $\mathcal{H}$ that only $\xi_{m+1} - \xi_m$ will enter into the calculation of the detection probability and not ξ_m and ξ_{m+1} separately. We call this quantity the *spacing* of the templates and denote it by δ. The whole template bank can now be constructed, using δ, as $\xi_k = \xi_{min} + k\delta$ $(k = 0, 1, \ldots)$ till ξ_{max} is reached.

We can now state a simple algorithm for setting up a one step search for given S_{min}, $Q_{d,min}$ and $Q_{0,max}$. A value for the spacing is chosen (starting from a large value, say $\delta = 0.10$ sec) and N_T is found. Then a threshold η is found, from Eq. (9), such that the false alarm becomes $Q_{0,max}$. The detection probability $Q_d(\eta;\, S_{min}, \xi_{min} + \delta/2)$ is found. If it exceeds $Q_{d,min}$, then stop or else reduce the spacing and repeat the above process.

3. Two step hierarchical search. We present below, a possible structure for a two step hierarchical search. It should be noted that this structure need not be unique. It is based on our experience with a one step search.

3.1. A two step hierarchical search: description. The basic idea behind a two step search is the use of two banks of templates. One of them has template chirp times placed farther apart than the other. A one step like search would be conducted with the finely spaced templates but only around "promising" candidate chirp times, namely, those templates in the coarser bank for which the maximum over their rectified output exceeds a threshold. This threshold would be kept lower than the one which would be used with the finely spaced templates.

We recall that the spacing for a one step search turned out to be a constant for the whole template bank because of the location invariance of the intrinsic ambiguity function. We expect the same feature for the two step template banks also. Thus we deal with two constant spacings in the following.

In a two step search as configured here, first the maximum over a rectified output is computed for each template in a bank B_1. We call B_1, the *first stage* template bank and denote the spacing of this bank by $\delta^{(1)}$. The number of templates used in B_1 would be $n_t^{(1)} = (\xi_{max} - \xi_{min})/\delta^{(1)}$. If for some template in B_1 having a chirp time ξ_m, it happens that the maximum over its rectified output, λ_m, crosses $\eta^{(1)}$ then we call this event a *crossing* of $\eta^{(1)}$ produced by the chirp time ξ_m. Given such a crossing, the next step involves using a template bank $B_{2,m}$ with a spacing $\delta^{(2)}$ that is smaller than $\delta^{(1)}$. We take $\delta^{(1)}/\delta^{(2)} = n$ to be an integer. This keeps the two banks of templates commensurate with each other.

The set of chirp times used in $B_{2,m}$ will be located symmetrically around ξ_m (except when $\xi_m = \xi_{min}$ or ξ_{max}, but these can be ignored). It will be convenient, therefore, to index the chirp times in $B_{2,m}$ with both positive and negative integers. Thus, the set of chirp times used in $B_{2,m}$ can be constructed as, $\xi_p = \xi_m + p\delta^{(2)}$. where $-n+1 \leq p \leq n-1$. However, the range of p need not be made as wide as this. For instance, it could be $-n/2 \leq p \leq n/2$ also. For the former, the number of template chirp times in $B_{2,m}$ would be $2(n-1)$ while for the latter it will be n.

Since $\eta^{(1)}$ would, in general, be kept quite low, the probability of more than one crossing in the first stage will not be negligible. In general, for every crossing of $\eta^{(1)}$, a fixed number, M $(n \leq M \leq 2(n-1))$, of templates will be employed as described above. Since the number of crossings that appear in our final results is small (typically, ~ 2), the choice of M within the above mentioned range does not make too much of a difference to the computational cost. We choose $M = 2(n-1)$ for our analysis, the maximum of the range, in order to maximize our chances of detection. Thus, the templates, in $B_{2,m}$, with chirp times ξ_{n-1} and ξ_{-n+1}, will have a separation of $\delta^{(2)}$ from the templates corresponding to ξ_{m+1} and ξ_{m-1} respectively.

Let n_c be the number of crossings that are produced by the first stage templates in B_1. Then the total number of *second stage* templates that will be used will be $n_c M$. Adjacent crossings will reduce this number since there would be some second stage templates in common for such crossings. It is easily seen, however, that adjacent crossings have a negligible probability compared to non-adjacent ones. Finally, the overall maximum over the rectified outputs of the second stage templates employed is found. We denote it by $\Lambda^{(2)}$. If $\Lambda^{(2)}$ crosses a threshold $\eta^{(2)}$ $(> \eta^{(1)})$, a detection is announced. Thus, $\Lambda^{(2)}$ is the test statistic for a two step search configured as above.

3.2. Determination of thresholds and spacings. We impose on the two step search described above, conditions similar to the one step template placement criteria of section 2.4. The *two step template placement criteria* are : **(i)** Every signal with a strength greater than a given minimum strength S_{min} should produce, with a probability $Q_{d,min}$, at least one crossing among the two templates which lie on either side of it. It should also be detected with a probability of $Q_{d,min}$ when the second stage templates corresponding to the above crossings are employed. **(ii)** The false alarm should be less than a specified level $Q_{0,max}$. This false alarm is for the overall search and does not refer to a specific level of the hierarchy. As in the case of a one step search, a solution in terms of thresholds and spacings need not exist for all combination of S_{min}, $Q_{d,min}$ and $Q_{0,max}$. Our choice of only two adjacent templates for the first stage crossing is justified because most of the extra crossings obtained, when four templates are used instead of two, would actually be spurious since they will not lead to the final detection of the signal in the second stage.

If a crossing of $\eta^{(2)}$ (which is quite large) were to be induced by noise alone, it would imply that the noise "resembles" the template waveform very closely. Therefore, as in the case of an actual signal, one can expect that such a noise realization would also induce a crossing of the first stage threshold in a nearby first stage template. This need not be true when the templates are far apart. However, for small spacings (~ 0.030 sec) it can be expected that the presence of a hierarchy will not be an impediment to a false alarm. We have checked this using a Monte Carlo simulation.

We denote the average value of n_c in the absence of a signal by n_c^{av}. The average computational requirement of a two step search can then be expressed in terms of the total number of templates used on the average,

$$n_t^{av} = n_c^{av} \times M + n_t^{(1)} \; . \tag{11}$$

In Eq. (11), we have neglected, as before, the probability of adjacent crossings in the absence of a signal. Under the assumption of statistical independence made above, n_c^{av} can be obtained as,

$$n_c^{av} = Q_0(\eta^{(1)}) \times n_t^{(1)} \tag{12}$$

where $Q_0(\eta^{(1)})$ is the probability of a crossing for a single template, $Q_0(\eta) = 1 - F_0(\eta)$. $Q_0(\eta^{(1)})$ behaves almost like a step-function in a narrow range of $\eta^{(1)}$. This is of crucial importance in the following.

From Eq. (11), we see that that in order to reduce the computational requirement, $n_t^{(1)}$ should be made small or, equivalently, the first stage spacing, $\delta^{(1)}$, should be made large. However, an increase in $\delta^{(1)}$ will lower the observed strength, S_{obs}, of a signal having a chirp time $\xi = (\xi_m + \xi_{m+1})/2$, for $\xi_m \in B_1$. This would imply a decrease in the probability of a crossing induced by such a signal in the first stage and hence, a violation of criterion (i) above. To avert this, $\eta^{(1)}$ would have to be lowered too. However, $F_0(z)$ has an almost step-function like behavior below a critical value of z. If, in the course of increasing $\delta^{(1)}$, $\eta^{(1)}$ became less than this critical value, the value of n_c^{av} would rise quite fast so much so that n_t^{av} would actually *increase* with an increase of $\delta^{(1)}$ beyond this point. Thus, there should exist a solution for the thresholds and spacings in a two step search, for which the computational requirement is minimized. This optimum solution can be found by a simple extension of the algorithm that was presented for a one step search.

The Algorithm :

(i) Given S_{min}, $Q_{d,min}$, $Q_{0,max}$ and the padding N_p, a one step template bank and threshold is set up using the algorithm presented in Sec. 2.4.

(ii) A trial value of $\delta^{(1)}$ is chosen as $\delta^{(1)} = j \times \delta^{(2)}$ where, $j \geq 2$ is an integer. For each trial value of $\delta^{(1)}$, $\eta^{(1)}$ is calculated so that $Q_d(\eta^{(1)}; S_{min}, \xi_{min} + \delta^{(1)}/2) = Q_{d,min}$. The average computational requirement, n_t^{av}, is then calculated using Eq. (12) and Eq. (11). The value of $\delta^{(1)}$ is increased by incrementing j, starting from a suitable initial value, until the minimum of n_t^{av} is reached.

For small spacings of $\delta^{(2)} \sim 0.030$ sec, the number of templates that will be required in the one step search constructed in step (i) above would be ~ 1000. Thus, the typical threshold, $\eta^{(2)}$, that would be required is ~ 7.9 for a false alarm that leads to one false event per year on the average. The observed strength required to attain a detection probability of 0.95 for such thresholds is ~ 8.6. If the detection of signals having a strength $\simeq 8.6$ is desired with the above probability, then it is clear that an almost continuous set of template chirp times would be required since otherwise the observed strength would become < 8.6. Of course, this would require an infinite amount of computing power.

Thus, there does not exist a solution to the template placement criteria for such a set of values for $Q_{d,min}$,$Q_{0,max}$ and $S_{min} \leq 8.6$. We call such a limiting value of S_{min} as the *minimum visible strength* for a one step search. In actual practice, we find that the minimum visible strength is a little higher at 8.75. This is because $\eta^{(2)}$ also increases as $\delta^{(2)}$ is reduced.

The padding, N_p, was kept fixed throughout the discussion above. This parameter of a two step search is, however, decisive in an estimation of the computational *power* required.

3.3. Computational power required for an on-line analysis. For the on-line detection of a signal, it is required that the processing of a given segment of data be completed within the time required to gather the next one [4]. For a given N_p, each time series will have $N = N_p + \xi_{max}/\Delta$ samples, where Δ is the sampling interval. If an FFT is used, this implies doing $6N \log_2 N$ floating point operations (flops) [15]. The correlations will be followed by $3N_p$ flops for the squaring and summation required for the calculation of a rectified output. The maximization over a single rectified output would involve, at most, N_p flops. Thus, the total number of flops required per template chirp time, n_{flop}, is, $n_{flop} = 6N \log_2 N + 4N_p$. The total number of flops required for the whole template bank on the average, N_{flop}, is therefore $N_{flop} = n_t^{av} n_{flop}$. We neglect the relatively small number of flops involved in the calculation of $\Lambda^{(2)}$. Thus, for an on-line implementation of a two step search, N_{flop} operations would have to be performed in $N_p\Delta$ sec. The average computational power required, C_{online}, is then,

$$C_{online} = \frac{N_{flop}}{N_p\Delta} \times 10^{-6} \text{ MFlops}, \tag{13}$$

where, "MFlops" stands for a million floating point operations per second.

An increase in N_p leads to an increase in $F_0(\eta)$ for a given threshold. Consequently, the number of false crossings in the first stage would also increase. Since n_t^{av} starts to rise when the number of crossings is $\gtrsim 1$, the minimum of n_t^{av} for a larger N_p will be achieved at a larger value of $\eta^{(1)}$. At the same time, the requirement that the probability of a crossing be $Q_{d,min}$ will force $\delta^{(1)}$ to a smaller value since the required observed strength would now be higher. The overall effect is an increase in n_t^{av} as well as N_{flop}. On the other hand, an increase in N_p will result in a longer time in which the required processing has to be completed and hence a lower computational power. Thus, given S_{min}, $Q_{d,min}$ and $Q_{0,max}$, there would exist an optimum N_p at which C_{online} is minimized.

We compute the value of C_{online}, as a function of N, for two different ranges of the chirp time. For each range, the minimum values of n_t^{av} is found for a few representative values of N_p, keeping S_{min} fixed. This process is then repeated for progressively lower values of S_{min} till $\delta^{(2)}$ becomes ~ 0.030 sec. We quote our results for such values of S_{min} (note that these values are not the minimum visible strengths). Table 1 contains the results for $\xi_{min} = 2.0$ sec and $\xi_{max} = 32.0$ sec (corresponding to a 1.2, 1.2 $M_\odot$ binary).

In this table, $S_{min} = 8.8$. In Table 2, $\xi_{max} = 138.0$ sec (0.5 , 0.5 $M_\odot$ binary) and $S_{min} = 9.0$.

In each table the minimum value of n_t^{av} is computed, using the algorithm presented in the previous section, for several values of N. The value of C_{online} is then found at each

Table 1

Minimum C_{online} as a function of N for : $S_{min} = 8.8$, $\xi_{min} = 2.0$ sec, $\xi_{max} = 32.0$ sec, $Q_{d,min} = 0.95$. $\eta^{(2)} = 7.92$, $\delta^{(2)} = 0.0325$ sec.

$N \times \Delta$(sec)	C_{online}(MFlops)	$\eta^{(1)}$	$\delta^{(1)}$(sec)	n_t^{av}	n_c^{av}
64.0	41.3 (392.9)	5.58	0.358	97	1
128.0	32.6 (279.6)	5.75	0.325	107	1
256.0	31.6 (253.7)	5.92	0.293	115	1
512.0	33.6 (249.3)	6.11	0.260	124	1
1024.0	36.7 (253.2)	6.11	0.260	134	1

Table 2

Minimum C_{online} as a function of N for : $S_{min} = 9.0$, $\xi_{min} = 2.0$ sec, $\xi_{max} = 138.0$ sec, $Q_{d,min} = 0.95$. $\eta^{(2)} = 8.10$, $\delta^{(2)} = 0.0335$ sec.

$N \times \Delta$(sec)	C_{online}(MFlops)	$\eta^{(1)}$	$\delta^{(1)}$(sec)	n_t^{av}	n_c^{av}
256.0	234.2 (2092.1)	5.84	0.335	455	3
512.0	172.9 (1400.8)	6.03	0.301	502	3
1024.0	167.1 (1245.5)	6.21	0.268	545	3
2048.0	175.2 (1211.5)	6.21	0.268	588	6

such minimum. We also list the corresponding values of $\eta^{(1)}$, $\delta^{(1)}$, n_t^{av} and n_c^{av} (the last two are rounded to the nearest whole number). It is easy to show that $\eta^{(2)}$, and hence $\delta^{(2)}$, is independent of N_p. Their values are presented in the captions of the tables. The numbers in parenthesis in the second column are the computing powers required for an on-line one step search. The values of N are chosen as powers of two because an FFT is most efficient at these values [15]. The value of $Q_{0,max}$ is always chosen to give an average of one false event per year.

We call the ratio of the computing power required for an on-line one step search to that required for an on-line two step search as the *computational advantage* of a two step search. From Table 1, the computational advantage at the minimum of C_{online} is 8.0. In Table 2, the corresponding number is 7.5. The number of crossings, n_c, will have a variance given by $n_t^{(1)}\, Q_0(\eta^{(1)})\, \left(1 - Q_0(\eta^{(1)})\right)$. For the entry from Table 1 considered above, the r.m.s. deviation in $n_c M$ will be $\simeq 12$. Thus, the value of $n_t^{av} \simeq 127$ and the computational advantage falls to $\simeq 7.2$. This is not a large change. Thus, a two step search offers a large reduction in the computing power required for an on-line detection while providing a useful combination of detection and false alarm probabilities.

For $\xi_{max} = 32.0$ sec and the values of $Q_{d,min}$ and $Q_{0,max}$ that were used above, it was noted earlier that $S_{min} \sim 8.6$ is the minimum strength which would be detectable. At this value of S_{min}, the computational requirement would become infinitely large since a template would be required for each value of the chirp time. We find that as this limiting strength is approached, the computational advantage of a two step search increases to a value of ≈ 13. However, the second stage spacing becomes quite small for such low values of S_{min} which implies that the statistical correlations among the rectified outputs can no longer be ignored. The formula used for the detection probability would therefore be very erroneous in such a case. A more careful analysis, taking statistical correlations into account, is needed when the value of S_{min} becomes close to the minimum visible strength.

The probability of adjacent crossings was neglected as compared to non-adjacent crossings. This is true when $n_c \ll n_t^{(1)}$. The values obtained for n_c^{av} clearly satisfy this condition. Note that $\delta^{(1)}$ is large enough for statistical correlations to be negligible and hence Eq. (12) to be valid.

4. Conclusions. We have investigated the performance of a two step hierarchical search for the detection of Newtonian waveforms from coalescing binaries. The noise power spectral density used in the analysis is that of the initial LIGO. A rigorous formalism to describe a two step search was presented which employs the detection probability of a signal in an essential way to set up the bank of templates and thresholds.

Our main result is that, as compared to a one step search, a two step search can reduce the computing power required for an on-line detection of Newtonian signals by at least a factor of $\simeq 8$. For an on-line detection of signals having a strength of ~ 8.8 (detection probability $\simeq 0.95$ and an average of one false event per year) the computing power required, for a two step search, is 167 MFlops when the range of chirp times is taken as $\xi_{min} = 2.0$ sec and $\xi_{max} = 138.0$ sec. We expect our results to hold good since the second stage spacings are ~ 0.030 sec. For weaker signals the spacings turn out to be much smaller, in which case statistical correlations will play a very significant role. The formula used for the detection probability would then be suspect. However, if we apply this formula for smaller spacings, the reduction achieved in computational power turns out to be much larger (a factor of ≈ 13). But these cases merit a more thorough investigation which should consider statistical correlations more carefully.

Acknowledgements. We are grateful to Prof. B. F. Schutz for pointing out an error in the calculation of the false alarm for a two step search. Special thanks are due from SDM to Dr. Soma Mukherjee for help with the statistical literature. SDM is supported by a CSIR fellowship.

References

[1] A. Abramovici *et al.*, Science, **256**, 325 (1992).
[2] C. Bradaschia *et al.*, Nucl. Inst. A., **289**, 518 (1990).
[3] K.S. Thorne, in *300 Years of Gravitation*, S.W. Hawking and W.Israel (eds.), (Cambridge Univ. Press, 1987).
[4] B.F. Schutz, in *The Detection of Gravitational Radiation,* edited by D. Blair (Cambridge, 1989) pp 406-427.
[5] C.W. Helstrom, *Statistical Theory of Signal Detection,* 2nd. ed, (Pergamon Press, London, 1968).
[6] B.S. Sathyaprakash and S.V. Dhurandhar, Phys. Rev. D **44**, 3819 (1991).

[7] S.V. Dhurandhar and B.S. Sathyaprakash, Phys. Rev. D **49**, 1707 (1994).
[8] B. Owen, submitted to Phys. Rev. D (Caltech preprint).
[9] T. A. Apostolatos, submitted to Phys. Rev. D (Preprint : Max-Planck, Jena).
[10] Workshop on Coalescing Binaries, *Conf. on Astrophysical Sources of Gravitational waves,* Pennsylvania State University (July 1995).
[11] J. P. A. Clarke and D. M. Eardley, Astrophys. J. **215**, 311, (1977).
[12] L.S. Finn and D.F. Chernoff, Phys. Rev. D **47**, 2198 (1993).
[13] Press, Flannery, Teukolsky, Vetterling, *Numerical Recipes* (Cambridge Univ. Press, 1986).
[14] C. Cutler and E. E. Flanagan, Phys. Rev. D **49**, 2658 (1994).
[15] O. E. Brigham, *Fast Fourier transform and its applications*, (Prentice-Hall, Englewood Cliffs, 1988).
[16] R. N. Bracewell, *Fourier transform and its applications*, 2nd ed. (McGraw-Hill, New York, 1986).
[17] J. W. Goodman, *Statistical Optics*, (John Wiley, New York, 1985).
[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).