

# Extended hierarchical search (EHS) algorithm for detection of gravitational waves from inspiralling compact binaries

Anand S Sengupta<sup>1</sup>, Sanjeev V Dhurandhar<sup>1</sup>, Albert Lazzarini<sup>2</sup>  
and Tom Prince<sup>3</sup>

<sup>1</sup> Inter-University Centre for Astronomy and Astrophysics, Post Bag 4, Ganeshkhind,  
Pune 411 007, India

<sup>2</sup> LIGO Laboratory, California Institute of Technology, Pasadena, CA 91125, USA

<sup>3</sup> Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA

E-mail: anandss@iucaa.ernet.in, sanjeev@iucaa.ernet.in, lazz@ligo.caltech.edu  
and prince@caltech.edu

Received 1 October 2001, in final form 14 October 2001

Published 14 March 2002

Online at [stacks.iop.org/CQG/19/1507](http://stacks.iop.org/CQG/19/1507)

## Abstract

Pattern matching techniques such as matched filtering will be used for online extraction of gravitational wave signals buried inside detector noise. This involves cross correlating the detector output with hundreds of thousands of templates spanning a multi-dimensional parameter space, which is very expensive computationally. A faster implementation algorithm was devised by Mohanty and Dhurandhar using a hierarchy of templates over the mass parameters, which speeded up the procedure by about 25–30 times. We show that a further reduction in computational cost is possible if we extend the hierarchy paradigm to an extra parameter, namely, the time of arrival of the signal. In the first stage, the chirp waveform is cut-off at a relatively low frequency allowing the data to be coarsely sampled leading to cost saving in performing the FFTs. This is possible because most of the signal power is at low frequencies, and therefore the advantage due to hierarchy over masses is not compromised. Results are obtained for spinless templates up to the second post-Newtonian (2PN) order for a single detector with LIGO I noise power spectral density. We estimate that the gain in computational cost over a flat search is about 100, which is about a factor of four over the previous less sophisticated hierarchical scheme.

PACS numbers: 0480N, 0705K, 9555Y, 9780

(Some figures in this article are in colour only in the electronic version)

## 1. Introduction

Massive compact binary systems consisting of neutron stars (up to a distance of 20 Mpc) or black holes (up to 200 Mpc) ranging in masses from  $\sim 0.5M_\odot$  to  $\sim 10.0M_\odot$  are one of the most promising sources for large scale laser interferometric detectors for gravitational waves (GW). A few minutes before the final merger, the GW frequency of such sources will lie within the bandwidth of the detectors. The expected GW waveform of these sources is known adequately to allow pattern matching techniques such as matched filtering to be used for signal extraction.

However, matched filtering is computationally quite expensive, since hundreds of thousands of templates must be used to search the multi-dimensional parameter space. The hierarchical search algorithm of Mohanty and Dhurandhar [1], which employed a hierarchy of templates over the mass parameters, achieved a significant improvement in cutting down the cost by a factor between 25 and 30. In this paper, we explore the possibility of extending the hierarchical paradigm to an extra dimension—the time of arrival of the signal. This is achieved by using the idea of decimating the signal in time (proposed before in [2] but not worked out in detail). This reduces the computational cost by about a factor of 100 over the one-step search and by a factor of about four over the hierarchical search in [3], for a single detector with LIGO I noise power spectral density.

## 2. The one-step search

In the stationary phase approximation, the Fourier transform  $\tilde{h}(f)$  of the spinless restricted second-order post-Newtonian (2PN) waveform up to a constant factor  $\mathcal{N}$ , is given by

$$\tilde{h}(f) = \mathcal{N} f^{-7/6} \exp i \left[ -\frac{\pi}{4} - \Phi_0 + \Psi(f; M, \eta, t_0) \right] \quad (1)$$

where  $M$  is the total mass of the binary system,  $\eta$  is the ratio of the reduced mass to the total mass,  $\Phi_0$  and  $t_0$  are respectively the initial phase and the initial time when the frequency  $f$  attains the fiducial value  $f_a$ . The factor  $\mathcal{N}$  depends on the masses,  $f_a$  and the distance to the binaries [4].

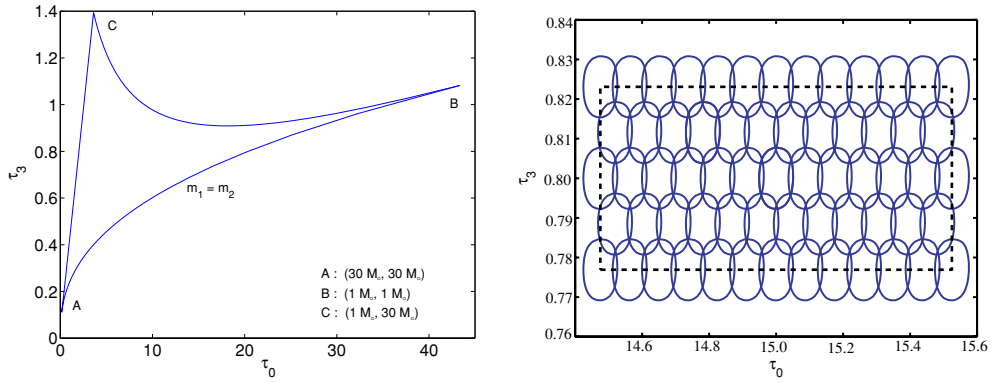
The function  $\Psi(f; M, \eta, t_0)$  describes the phase evolution of the inspiral waveform and is given by

$$\begin{aligned} \Psi(f; M, \eta, t_0) = & 2\pi f t_0 + \frac{3}{128\eta} \left[ (\pi M f)^{-5/3} + \left( \frac{3715}{756} + \frac{55}{9}\eta \right) (\pi M f)^{-1} - 16\pi (\pi M f)^{-2/3} \right. \\ & \left. + \left( \frac{15293365}{508032} + \frac{27145}{504}\eta + \frac{3085}{72}\eta^2 \right) (\pi M f)^{-1/3} \right]. \end{aligned} \quad (2)$$

The kinematical parameters ( $t_0, \Phi_0$ ) do not determine the shape of the waveform and can be treated very simply in the matched filtering procedure [5]. In what follows, the parameter space will explicitly refer to the space of dynamical parameters that determine the phase, and thus the shape of the waveform. For the 2PN case, it is the two-dimensional space described by  $M$  and  $\eta$ . In order to facilitate the analysis, one chooses a new set of parameters  $\tau_0$  and  $\tau_3$  in which the metric defined over the parameter space [6] is approximately constant. They are related to  $M$  and  $\eta$  by

$$\tau_0 = \frac{5}{256\pi\eta f_a} (\pi M f_a)^{-5/3} \quad \tau_3 = \frac{1}{8\eta f_a} (\pi M f_a)^{-2/3}. \quad (3)$$

Here  $\tau_0$  is the Newtonian time of coalescence and the chirp time  $\tau_3$  is related to the 1.5 PN order. In these new coordinates the parameter space is wedge shaped as shown in the left panel of figure 1.

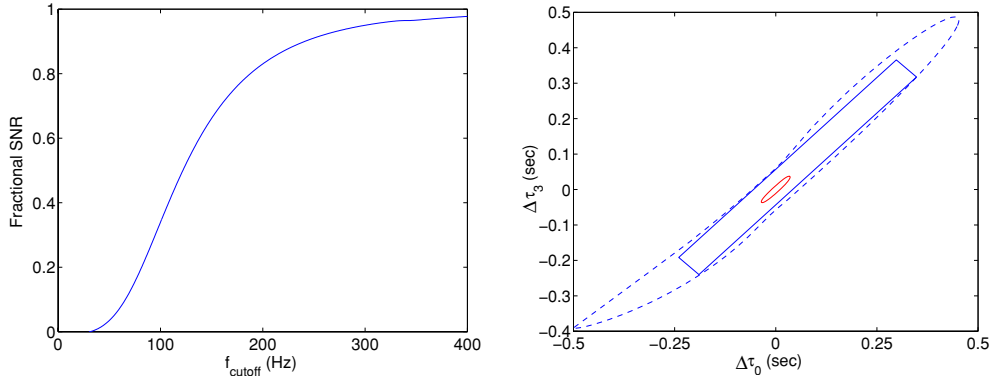


**Figure 1.** The left-hand panel shows the parameter space for 2PN waveforms in  $(\tau_0, \tau_3)$  space for binary masses in the range  $1.0M_\odot \leq m_1, m_2 \leq 30.0M_\odot$ . The fiducial frequency is set to  $f_a = 40.0$  Hz for this figure and the area is  $8.5 \text{ s}^2$ . Shown in the right-hand panel is a schematic of the discrete template placement using hexagonal closed packing (hcp) in a rectangle centred around  $\tau_0 = 15$  s,  $\tau_3 = 0.8$  s.

A straightforward implementation of matched filtering involves constructing a discrete lattice of normalized templates over the parameter space and computing correlations with the detector output. For our analysis, we take the LIGO laboratory accepted noise PSD<sup>4</sup> estimates for the 4 km initial configuration. Because the interferometer noise curves depend strongly on frequency, the sensitivity of the interferometers is practically limited to a bandwidth of hundreds of Hertz that encompasses the frequency of best sensitivity  $f_{\min} \approx 160$  Hz. It is sufficient to consider signals band limited by the seismic wall at  $f_s = 30$  Hz and the upper cut-off frequency  $f_u = 800$  Hz. The SNR lost due to the band limiting is  $\sim 0.14\%$ , which is negligible. The maximum of the correlation output is compared with the threshold  $\eta_2$  (we put the subscript ‘2’ because  $\eta_2$  will be used as the second stage threshold in the hierarchical scheme described later) which depends on a predecided false alarm rate usually set at one per year. The false alarm rate we assume here is for one detector only. If the data are sampled at 2048 Hz, the noise assumed to be a Gaussian stationary process and the number of templates  $\sim 10^4$  then the threshold  $\eta_2 \simeq 8.2$ . Detection is announced if the maximum correlation exceeds this threshold.

The degree of discretization of templates is governed by a maximum mismatch between the signal and the filter, which is taken to be 3% (corresponding to a maximum loss of event rate of 10%). The mismatch in terms of the parameters  $(\tau_0, \tau_3)$  is encoded in the ambiguity function  $\mathcal{H}(\Delta\tau_0, \Delta\tau_3)$ . Actually  $\mathcal{H}$  also depends weakly on the position  $(\tau_0, \tau_3)$ , but we ignore its effect in the present analysis. We use the most efficient hexagonal closed packed (hcp) tiling for placing the templates as shown in figure 1 (right panel), with a packing fraction of 83%. This leads to 17 000 templates for LIGO I noise PSD and for a mass range of 1.0–30.0  $M_\odot$ . Sampling at 2048 Hz with a data train 512 s long to allow for 75% padding, the number of points in the data train is  $2^{20}$ . The online speed required to cover the cost of FFTs amounts to 2.54 G-flops using this one-step implementation of matched filtering. The required speed rises to about 15 G-flops if the minimum mass of the mass range is reduced to 0.5  $M_\odot$  and to about 150 G-flops when the minimum mass is 0.2  $M_\odot$ .

<sup>4</sup> Plots and data files of LIGO I noise curve is maintained by J K Blackburn and is available at the URL [www.ligo.caltech.edu/~kent/ASIS\\_NM/initial.html](http://www.ligo.caltech.edu/~kent/ASIS_NM/initial.html).



**Figure 2.** The left panel shows the fraction of signal to noise ratio (SNR) extracted as a function of the cut-off frequency  $f_{\text{cut-off}}$ . The SNR is normalized to 1 for  $f_{\text{cut-off}} = f_u$ . The right panel shows the  $\mathcal{H}(\Delta\tau_0, \Delta\tau_3) = 0.80$  contour as a dashed line centred around  $\tau_0 = 15.0$  s and  $\tau_3 = 0.8$  s, along with the largest inscribed rectangle used in the trigger stage of the EHS algorithm for  $f_{\text{cut-off}} = 256$  Hz. The ellipse inside the rectangle is that used by the one-step search to tile the space and is given for comparison of the areas.

### 3. Extended hierarchical search

The hierarchical scheme as given in [1] involves parsing the data through two template banks instead of one. The basic idea is to use a coarse bank of templates in the first stage along with a lower threshold. This allows the parameter space to be scanned with a smaller number of templates. However, the lower threshold leads to a large number of false alarms which must then be followed up with a fine bank (maximum mismatch 3%) search.

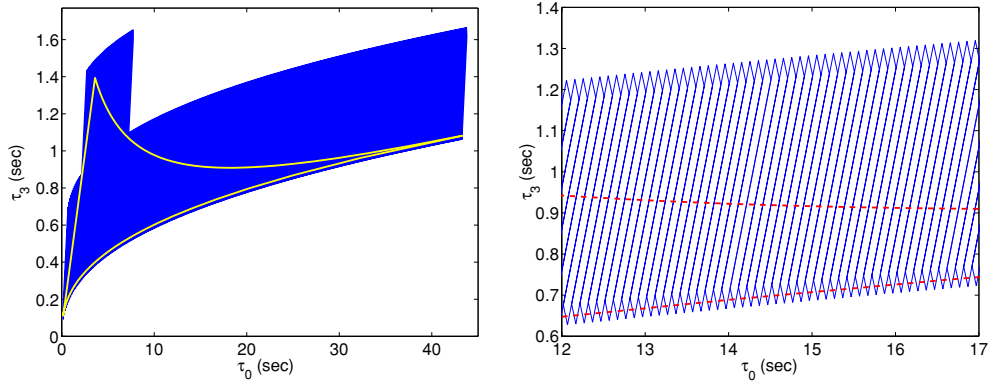
#### 3.1. Methodology

The above idea can be extended to include the time of arrival of the signal—by cutting off the chirp at a lower frequency  $f_{\text{cut-off}}$  and re-sampling it at a lower Nyquist rate in the trigger phase of a two-step implementation of matched filtering. This is possible because the inspiral signal contains most of the signal power at lower frequencies (the power spectrum scales as  $|\tilde{h}(f)|^2 \propto f^{-7/3}$ ) and also because of the nature of the colour of the LIGO I noise PSD, we do not lose much of the SNR. The fractional SNR that can be recovered in this case is given by  $\text{SNR}(f_{\text{cut-off}}) = I(f_{\text{cut-off}})/I(f_u)$ , where

$$I(f_{\text{cut-off}}) = 2 \int_{f_s}^{f_{\text{cut-off}}} \frac{df}{f^{7/3} S_h(f)} \quad f_s < f_{\text{cut-off}} \leq f_u. \quad (4)$$

In the above equation, we have ignored the plunge cut-off, which would apply for total masses above  $\sim 5M_\odot$  for the second stage and above about  $\sim 16M_\odot$  for the first stage. However, for these high mass ranges, the parameter space area is quite small, so that the overall gain factor should not be too much different. In the left panel of figure 2 we have plotted the relative SNR as a function of the cut-off frequency. As can be clearly seen, a significant fraction of the signal power can be recovered from a relatively small value of  $f_{\text{cut-off}}$ . Specifically, for  $f_{\text{cut-off}} = 256$  Hz, almost 92% of the SNR can be recovered. Since not too much of the SNR is lost, the hierarchy over the masses [3] is not compromised too severely.

The most obvious advantage of lowering the Nyquist rate in the first stage is that we have fewer points to contend with in computing the FFTs (the cost of FFTs scales as  $N \log_2 N$ ),



**Figure 3.** The left panel shows a possible schematic lay-out of first stage templates for implementing the EHS algorithm. Note that far too many templates have to be used in the first stage to cover the parameter space, due to inefficient tiling. The effect is most severe towards the tapering low-mass end of the parameter space. The right panel shows the details of first stage template placement at a higher resolution. The dashed line in this figure is the parameter space boundary.

leading to reduction in cost. For example,  $f_{\text{cut-off}} = 256$  Hz reduces the sampling rate to a quarter, reducing the cost of FFTs by almost the same factor. Secondly, the ambiguity function becomes wider thereby reducing the number of templates used in the first stage and reducing the cost.

The first stage threshold  $\eta_1$  is set by striking a balance between the two opposing effects; (i)  $\eta_1$  must be low enough to allow coarsest possible placement of trigger stage templates so that the first stage computational cost is minimized, (ii)  $\eta_1$  must be high enough to reduce the number of false crossings, so that the second stage cost is reduced.

By choosing  $\eta_1$  optimally, the total computational cost arising from both stages is minimized.

### 3.2. Computational cost of the EHS

We consider data trains of length 512 s which accommodate chirps of maximum length 95 s with sufficient padding corresponding to a minimum mass limit of  $1.0M_{\odot}$ . We consider first and second stage sampling frequencies of 512 Hz and 2048 Hz, respectively. The cut-off frequency in the first stage is set at  $f_{\text{cut-off}} = 256$  Hz. Thus the data points to be processed are  $2^{18}$  and  $2^{20}$  for the coarse and fine banks, respectively. The second stage threshold is set at  $\eta_2 \simeq 8.2$  as before (see section 2) for a false alarm rate of one per year. The signals of minimal strength that can be observed, assuming a minimum detection probability of 95% and 3% mismatch between templates, turns out to be 9.17. The first stage SNR for the above-mentioned cut-off can be calculated to be 92% of 9.17, which is 8.44. It is found that the total cost is minimized for  $\eta_1 \approx 6$ . Again, assuming a minimum detection probability of 95% in the first stage itself, the ambiguity function can be allowed to drop to 6.75 which is about 80% of its maximum value. This large drop allows for very big tiles as shown in the right panel of figure 2. However, the awkward shape of these tiles makes the tiling parameter space a difficult proposition. We surmount this problem by cutting out the largest inscribed rectangle as shown in the same figure, which has an area of  $0.05 \text{ s}^2$ . This is  $\sim 50$  times larger than the area of the second-stage tiles. Combining the first and second stage costs, the required online

speed is then estimated at 19.54 M-flops, giving a gain factor of about 130 over the one-step search.

#### 4. Discussion

We have shown that a three-dimensional hierarchical search for gravitational waves emitted by inspiralling compact binaries is a highly promising proposition. We have obtained a computational cost reduction factor of about 130 for an ideal search, when the parameter space can be optimally tiled by the first stage template rectangles. However, due to *boundary effects* this factor can be drastically reduced. As can be seen in figure 3, the parameter space tapers towards the low-mass end where the tiling becomes inefficient. Most of the tiles go ‘out’ of the deemed parameter space. It is estimated that this can cut down the above-mentioned gain factor by about 50%. These effects become less pronounced when the lower mass limit of the mass range is reduced. We expect to recover the gain factor to about 100, if the lower mass limit is reduced to less than  $0.5M_{\odot}$ .

The second important effect affecting the gain factor will come from the rotation of the first and second stage tiles and also the differential rotation between them. These effects have still to be taken into account and will be the thrust of our future investigations. However, we do not expect this effect to alter the gain factor very much.

#### Acknowledgments

The authors would like to thank A Vicere, B S Sathyaprakash and S Mohanty for fruitful discussions. SVD would like to thank B Barish and G Sanders for a visit to Caltech. AS would like to thank CSIR for JRF. This work was partially supported by the LIGO Laboratory under NSF cooperative agreement PHY92-10038. This document has been assigned LIGO Laboratory document number LIGO-P010020-00-E.

#### References

- [1] Mohanty S D and Dhurandhar S V 1996 *Phys. Rev. D* **54** 7108
- [2] Tanaka T and Tagoshi T 2000 *Phys. Rev. D* **62** 082001
- [3] Mohanty S D 1998 *Phys. Rev. D* **57** 630
- [4] Dhurandhar S V and Sathyaprakash B S 1994 *Phys. Rev. D* **49** 1707
- [5] Sathyaprakash B S and Dhurandhar S V 1991 *Phys. Rev. D* **44** 3819
- [6] Owen B J 1996 *Phys. Rev. D* **57** 6749