

Building a stochastic template bank for detecting massive black hole binaries

Stanislav Babak

Max-Planck-Institut für Gravitationsphysik, Albert-Einstein-Institut, Am Mühlenberg 1,
D-14476 Golm bei Potsdam, Germany

E-mail: stba@aei.mpg.de

Received 4 April 2008, in final form 1 July 2008

Published 16 September 2008

Online at stacks.iop.org/CQG/25/195011

Abstract

The coalescence of pairs of massive black holes are the strongest and most promising sources for LISA. In fact, the gravitational wave signal from the final inspiral and merger will be detectable throughout the universe. In this paper we describe the first step in a two-step hierarchical search for the gravitational wave signal from inspiraling massive BH binaries. It is based on a method routinely used in ground-based gravitational wave astronomy, namely filtering the data through a bank of templates. However we use a novel, Monte Carlo based (stochastic), method for laying a grid in the parameter space, and we use the likelihood maximized analytically over some parameters, known as the \mathcal{F} -statistic, as a detection statistic. We build a coarse template bank to detect gravitational wave signals and to make preliminary parameter estimation. The best candidates will be followed up using a Metropolis–Hasting stochastic search to refine the parameter estimates. We demonstrate the performance of the method by applying it to the Mock LISA data challenge 1B (training data set).

PACS numbers: 04.30.Db, 04.80.Nn

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

1. Introduction

Current astrophysical observations show abundant evidence that almost all galaxies contain (super)massive black holes (MBH) in their nuclei. This was proposed as early as the 1960s [1] to explain the enormous luminosity of quasars through accretion of gas onto the MBH. The common belief is that a large fraction of galaxies must have been quasars in the past. The study of the kinematics of stars and/or gas in the central galactic regions suggests the existence of a dark compact object with a mass in the range 10^6 – $10^9 M_{\odot}$, which is most probably a MBH

[2]. The accretion mechanism cannot explain the growth in mass of all BHs from the quasar stage to the present day, which suggests that they were formed by merging smaller mass BHs [3], and, in fact, we do observe many merging galaxies. The most famous is a direct VLBA observation suggesting the existence of two MBHs with a total mass $10^8 M_\odot$ and an orbital separation ~ 7 pc in the center of radio galaxy 0402+379 [4].

The last week of the inspiral of a MBH binary followed by the merger to form a single fast spinning excited BH is the most powerful source of gravitational waves (GWs). In fact, we will be able to detect the events throughout the universe with LISA. LISA [5] is a proposed space-borne gravitational wave observatory which will be launched in 2018+ and will be sensitive to GW signals in the range 10^{-4} –0.1 Hz. It will consist of three identical spacecraft forming an equilateral triangle on a heliocentric orbit.

In this paper, we will concentrate only on the detection of the inspiralling part of the GW from the coalescing MBH binary. This part is quite well modeled analytically using the post-Newtonian (PN) approach, which is an iterative solution of the Einstein equations for the two-body problem using an expansion in the orbital velocity, v ,¹ assumed to be small. We restrict our attention here to the inspiral of non-spinning BHs moving on a quasi-circular trajectory. Furthermore, we use restricted PN waveforms (keeping only a leading order amplitude of the GW signal) with the phase defined up to 2nd PN order [6]. These are not the most accurate waveforms available today, but the purpose of the current work is to describe a data analysis algorithm and to demonstrate its performance, and this work can be extended to more sophisticated models in the future. Our conventions for the LISA model are the same as adopted for the mock LISA data challenge (MLDC) [7, 8].

This paper describes the first step in a two-step hierarchical search method. It is based on constructing a coarse grid of templates and (matched) filtering the data through each template. The candidates will be taken as input to the second stage of the search which will utilize a simplified Metropolis–Hasting stochastic search. Before we begin a detailed description of our method we should mention the alternatives.

The most successful method so far is based entirely on the Metropolis–Hastings stochastic search [9–11]. We cannot call it a Markov chain as the method violates the Markovian property (it uses previous information to inform some proposal distributions for further jumps). The success of this method lies in using a range of proposal distributions that utilize symmetries in the detector’s response function and global properties of the likelihood surface. This, in combination with simulated and frequency annealing schemes, makes this method very fast (can be run on a laptop) and robust. However in the presence of several sources one has to find signals one by one: finding a signal, estimating its parameters, removing it and searching for the next one. This somewhat diminishes the efficiency of this method.

Another method [12] is closer to that proposed in this paper and employs three stages: (i) construct a time–frequency map of the data, identify the chirp and perform χ^2 fitting to estimate the masses and coalescence time (ii) use a template grid (based on [13]) to do a matched filtering search and (iii) use the Metropolis–Hastings stochastic search to follow up on the candidates from the previous steps. This search method has a lot of similarities to what we will describe below, but our implementation is very different. The advantage of using multi-stage methods as described above and presented in this paper is the estimation of the number of signals present in the data. By mapping the likelihood throughout the whole parameter space using a template grid we get straight away the set of all candidates (triggers, clustered in time of coalescence, which crossed a preset signal-to-noise ratio (SNR) threshold).

¹ In fact it is an expansion in v/c , but throughout this paper we shall work in geometrical units $G = c = 1$.

Now we are ready to discuss our method. The main novelty is the new template bank: we have used a stochastic method for building it and this is the *first* implementation of such a method. The focus of this paper will be on the construction of this template bank. The main idea is simple: you throw templates in the parameter space and compute the match (also called overlap) between the trial template and other templates in the bank, the trial template is then added to the bank if the maximum match with the other templates is below a preset minimum. The overlap is defined as a Hermitian inner product between two normalized signals:

$$(\hat{s}|\hat{h}) \equiv 2 \int_0^\infty \frac{\tilde{s}^*(f)\tilde{h}(f) + \tilde{s}(f)\tilde{h}^*(f)}{S_n(f)} df. \quad (1.1)$$

Here $S_n(f)$ is the one-sided noise power spectral density, \tilde{s} and \tilde{h} are signals in the Fourier domain, normalized such that $(\hat{s}|\hat{s}) = (\hat{h}|\hat{h}) = 1$.

The template bank can be seen as a grid in the parameter space and the coarseness of the grid is defined by the minimal match (MM) between nearby templates (proper distance between the neighboring points on the grid). For ground-based interferometers MM is chosen as a compromise between computational power and a tolerated loss in detection volume. In the case of LISA, the signals are usually strong and detection will not be a problem, so our choice of MM is governed by the speed of computation and by the closeness of the estimated parameters to the true values. Given MM , we want to construct the template bank optimally: minimize the number of templates while trying to keep the overlap between neighboring templates close to MM . The problem of template placement was attacked in the gravitational wave community by several groups [13–16]. The last two works also deal with the stochastic placement of the template grid, but use different implementation. Previous banks aimed at a (sub)optimal placement using the metric induced on the parameter space. The main problem encountered in those methods is to find a coordinate frame which is close to Cartesian if the parameter manifold is flat (or almost flat) and which makes the ambiguity contours² close to circles. Even if one overcomes this problem, there is still a need to deal with boundaries which are usually overpopulated with templates. In addition, an attempt to cover the parameter space completely (without holes) inevitably forces us to place templates with overlapping ambiguity contours, which leads to an unnecessarily large number of templates. The last problem was also partially solved in [16] where the authors use different implementation of the stochastic method. We will allow for small holes in covering the parameter space but increase slightly MM .

There could be several different ways to construct a stochastic bank in practice. We describe one way of building such a bank in section 2, and this is the main result of this paper. Then we use the constructed bank to perform filtering of the noiseless data; these results are presented in section 3. Here we consider a simplified signal from a MBH binary, which is the restricted 2PN model adopted in MLDC [17] and summarized in appendix A. The signal depends on five intrinsic parameters: the masses (m_1, m_2), time of coalescence (T_c) and sky positions (θ_s, ϕ_s) and four extrinsic parameters. In our search we have used the \mathcal{F} -statistic [18] to maximize over the extrinsic parameters: inclination (ι), polarization (ψ), luminosity distance (D_L) and initial orbital phase (ϕ_0), and we use an iterative correlation method for identifying the time of coalescence. The effect of the coalescence time on the ambiguity contours is discussed in appendix B. We summarize the key points of this paper in section 4.

² Ambiguity contours are contours of constant overlap in the parameter space.

2. Building a stochastic bank

Let us repeat the main idea behind building the stochastic template bank. We start with choosing a random point in the parameter space and this is our first template in the bank. Then we loop over N trials and for each trial (a randomly chosen point in the parameter space within the prior) we compute the overlaps between the trial template and each template already in the bank. If the maximum overlap is below MM we add the point to the template bank, and reject it otherwise. If MM is high, say above 0.95 one can usually use the metric induced on the parameter space to compute the distance between the trial template and the neighboring templates in the bank:

$$ds^2 = g_{ij} \delta\lambda_i \delta\lambda_j. \quad (2.1)$$

Here g_{ij} is the metric

$$g_{ij} = \left(\frac{d\hat{h}}{d\lambda_i} \middle| \frac{d\hat{h}}{d\lambda_j} \right), \quad (2.2)$$

where we have assumed that the waveform $h(t, \lambda_i)$ depends on the parameters λ_i and $\delta\lambda_i$ are the mismatches in the parameters. For closely placed templates, the overlap, \mathcal{O} , can be approximated as

$$\mathcal{O} = 1 - ds^2; \quad (2.3)$$

however if we increase the coarseness of the grid the quadratic approximation is not sufficient and using this approximation usually overestimates the overlap. Moreover, the metric gives only a local measure of distances between templates, whereas the overlap is a global measure and it is important for placing templates, as we will show later in this section. For our search we chose the minimal match to be 0.88 and this is too low to use the metric to estimate the overlap.

Computing the overlap directly is a computationally expensive procedure. To tackle this problem we use two tricks. First, we assume that the coalescence time is 5 h after the end of the observation period which is taken to be 1 year. The duration of the observation is motivated to include one full cycle in LISA motion and therefore in the Doppler modulation of the signal. The choice of the coalescence time was mainly governed by our desire to have the highest frequency of the templates below 1 mHz, which allows us to use the long wavelength approximation to the LISA response. We have only used the X TDI channel in the unequal arm Michelson combination (see appendix A for more details on the signal model). Restricting the highest frequency of the signal also enabled us to use a low sampling rate. In the end, we were able to compute about 200 overlaps per second on a single core of an Intel DUO 2 computer. The second trick was to estimate the density of templates in the parameter space, then split it into regions of approximately equal (small) number of templates, and to populate each sub-region separately using different nodes on a cluster of computers. The small number of templates allows us to do many trials and have a reasonable overall coverage. Having about equal expected number of templates in each sub-region balances the computational load on the different processors.

We have noted that for fixed masses the ambiguity contours which correspond to $\mathcal{O} = 0.88$ do not change significantly for different sky locations. This implies that we can use a fixed sky position, lay a template grid in m_1 – m_2 space and it will be a valid grid (with good approximation) for any sky position. Moreover $MM = 0.88$ is too low for placing templates in the sky position, as the ambiguity contours are very big, and we use instead a uniform grid on the sphere with approximately 2700 points. The separation of points on the sphere

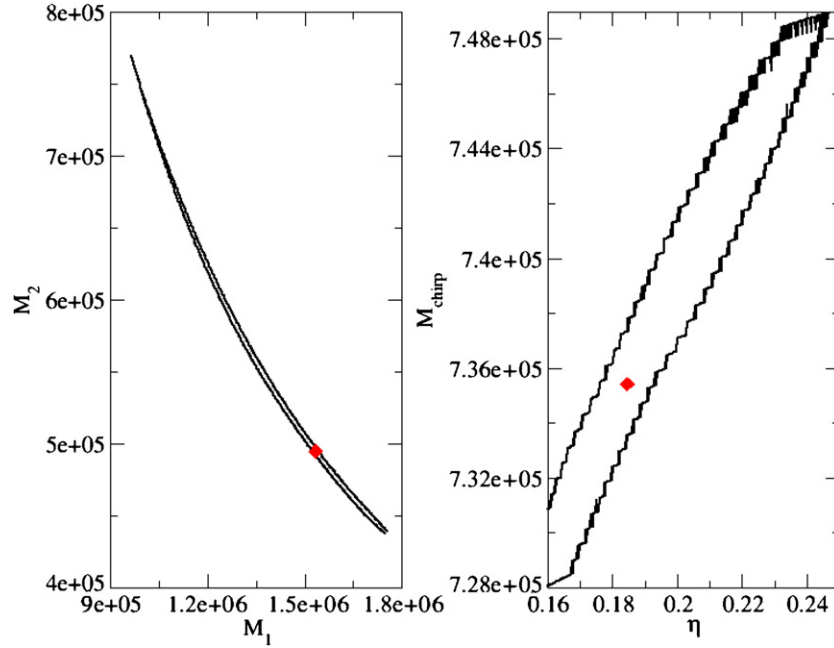


Figure 1. Global ambiguity contours: left panel in m_1 – m_2 and right panel in η – \mathcal{M} coordinates. The central point is depicted as a diamond on both plots.

was taken as double the expected error in the determination of the sky location as defined by the inverse Fisher matrix. We have also found that the other (extrinsic) parameters do not affect the ambiguity contours, meaning that the contours stayed almost the same over several arbitrary chosen values, so we used a particular choice ($\iota = 0$, $\psi = 0$, $\phi_0 = 0$). Consequently, we had to place templates in the m_1 – m_2 plane only. To search for coalescence time we slide the templates across the data and look for the maximum of correlation. We have found that the ambiguity contours are rather sensitive to the particular (fixed) choice of coalescence time, and they change not only in size but also orientation. The chosen coalescence time is not the most conservative, but this is cured by an iterative maximization of the \mathcal{F} -statistic over this parameter during the search. One can find more details on this in the following section and in appendix B.

As mentioned above, we first need to estimate the density of templates in m_1 – m_2 . We have found that m_1 , m_2 are not convenient coordinates. In figure 1 we show the typical *global* ambiguity contours of match 0.88 (we will explain the meaning of ‘global’ a bit later) in m_1 – m_2 space and in η – \mathcal{M} space. Here masses are red-shifted masses in units of solar mass, $\eta = m_1 m_2 / M^2$ —the symmetric mass ratio, $\mathcal{M} = M \eta^{3/5}$ is the chirp mass and $M = m_1 + m_2$. One can see that the ambiguity contour in masses is almost a one-dimensional structure, and in the \mathcal{M} – η space the chirp mass changes by less than 3% over the whole range of η . We should mention that we have restricted the parameter space to the prior adopted in the MLDC: $m_1 \in [10^6, 5 \times 10^6] M_\odot$, $m_1/m_2 \in [1, 4]$. The size of the contours does change across the parameter space, and they also rotate, but the angle of inclination to the η -axis is always small. This makes perfect sense, because the first, dominant, term in the gravitational wave phase depends only on \mathcal{M} . These arguments are convincing enough to adopt η , \mathcal{M} as the coordinate frame for estimating the density of templates.

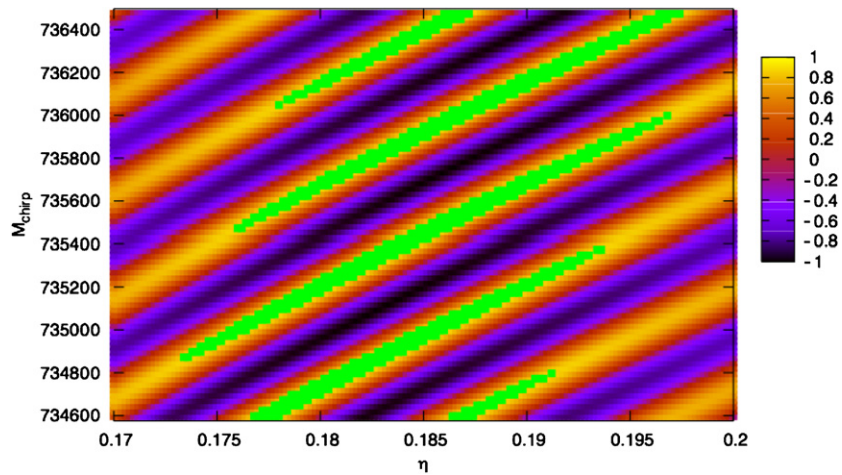


Figure 2. Internal structure inside the global ambiguity contour. This is a 3D plot projected onto the η - \mathcal{M} plane and the z -axis (overlap with the central point) is color-coded. The overlapped green points (gray shaded ellipses in black-and-white version) are those with overlap above $MM = 0.88$.

Before we go further we need to explain the meaning of ‘global’ ambiguity contours. Inside those global contours the overlap exhibits damping oscillations, with maximum amplitude of one at the central point and damps to $MM = 0.88$ at the boundaries of the global contour. In figure 2 we zoom into a part of the global contour around the central point ($\eta = 0.1845$, $\mathcal{M} = 735423M_{\odot}$). The overlap with the central point is color-coded, and the green points correspond to an overlap above MM . One can see separated maxima within the global contour. This is not something new, since such a structure was first described in [19], and later in the context of LISA in [10]. The size, orientation and height (maximum overlap) of each small ellipse within the global contour as well as their separation depends on m_1 , m_2 and T_c .

Let us emphasize an important point; the ambiguity ellipse defined by MM and described locally by the metric on the parameter space corresponds only to the green ellipse around the central point in figure 2, and ignores other maxima. Taking only the central small ellipse as an ‘area’ of the template in the parameter space will seriously overpopulate the parameter space with templates (even if the placement is nearly optimal as in [16]). Here we are using the *global* properties of the parameter space by computing overlap. The secondary maxima, being well separated from the primary, increase significantly the area of the parameter space covered by a single template. This helps to reduce the size of the bank, and here we see a big advantage of using the overlap over the metric-based methods for the template placement. We will present a more detailed study of the global properties of the likelihood using \mathcal{F} -statistic (rather than overlap) in the separate publication.

The distribution of the overlap inside the contour is presented in figure 3. Note that the orientation of the global ambiguity contour is different from the orientation of the local ellipses within it.

We now return back to the issue of estimation of the template density. We will define the area of a template as the area inside the global ambiguity contour. This is clearly an overestimate of the area of a template, since there are ‘holes’ inside the global contour, but at this stage we just want to split the parameter space into regions containing approximately equal number of templates, and are not doing rigorous template counting. We compute the

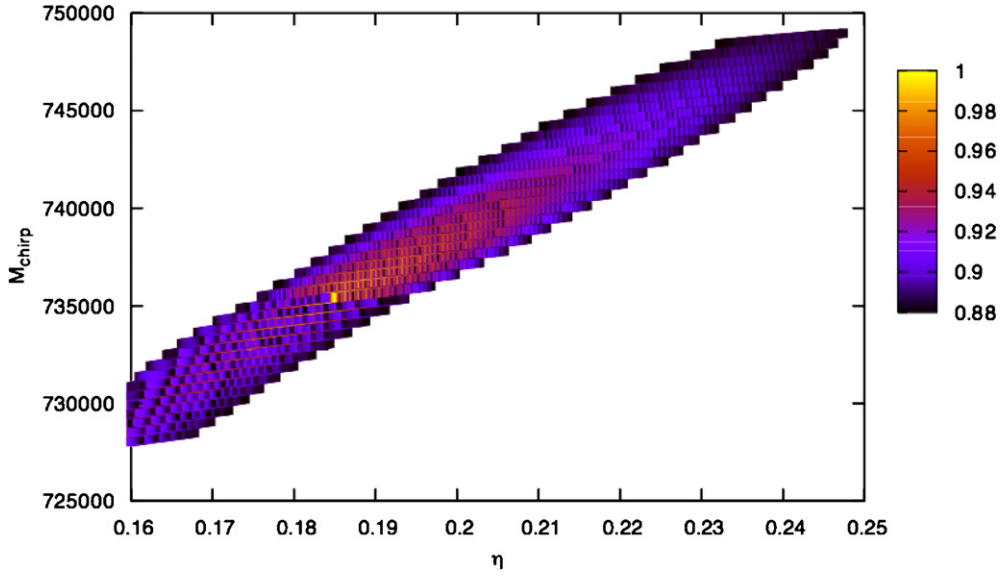


Figure 3. Distribution of the overlap inside the global ambiguity contour. It is a 3D plot projected onto the η - \mathcal{M} plane and the z -axis (overlap with the central point) is color-coded.

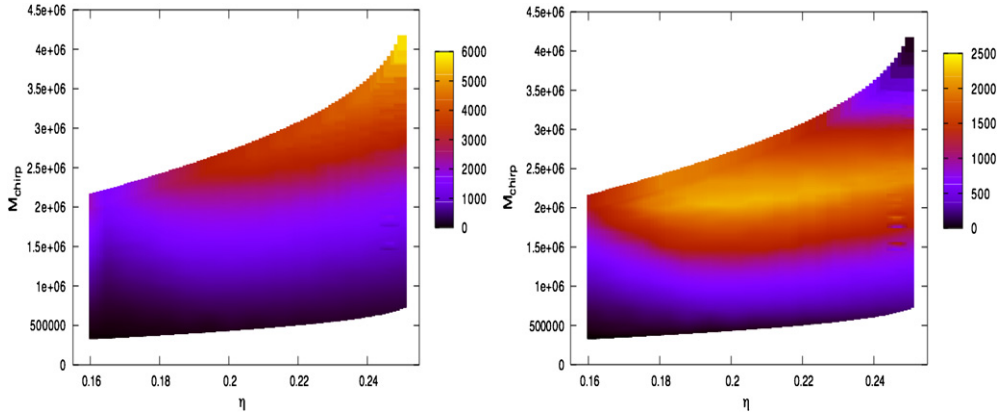


Figure 4. Interpolated area of a template. Left panel: area of template as defined by the global ambiguity contour across parameter space; right panel: useful area of a template (the part of the global ambiguity contour that lies within the inside parameter range defined by the priors).

area of 336 templates, uniformly distributed over the parameter space. Then we compute number of templates as in [20], assuming that $\sqrt{\|\det g_{ij}\|} \sim (1 - MM)/A_t$, where A_t is the interpolated (coordinate-dependent) area of a template:

$$N \sim \frac{\int \sqrt{\|\det g_{ij}\|} dx dy}{(1 - MM)} \approx \int \frac{dx dy}{A_t}. \quad (2.4)$$

Again, the absolute number is not very important, but what is important is how it changes across parameter space. In figure 4, we plot the interpolated total area of the templates (in the left panel) and the area of the templates that falls inside the parameter region defined by

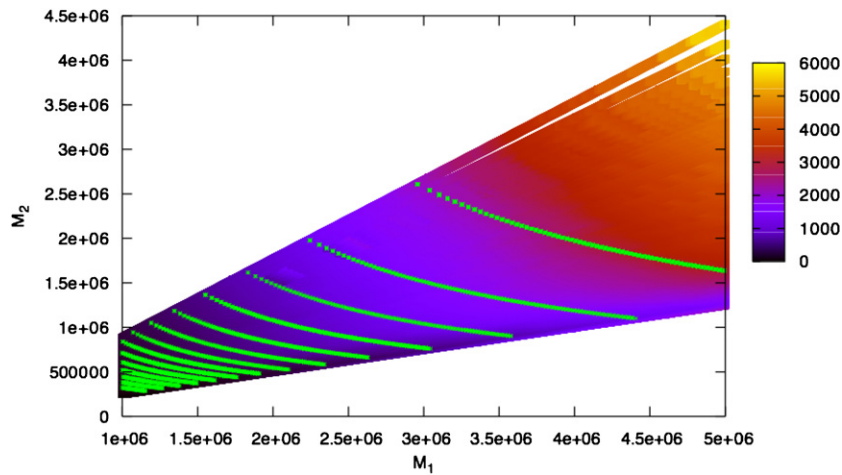


Figure 5. Splitting of the $m_1 - m_2$ parameter space into 15 sub-regions. Green crosses correspond to the region boundaries, and the color-coded z -axis gives the area covered by each template. (This figure is in colour only in the electronic version)

the priors (in the right panel). One can see that (i) contours of equal area again correspond approximately to levels of constant \mathcal{M} ; (ii) splitting the parameter space using the areas in the left or right panels is not the same. We have used the left panel (total area of the template) to do the splitting of the parameter space, because we are using the assumption that the area has an elliptical shape as in equation (2.4), which is a better approximation for the total area than the area within the prior. However we will keep in mind that the expected deviation in the number of templates from the mean will be higher in the $\{0.16, 5.e5\}$ and the $\{0.25, 4.e6\}$ corners of the $\eta - \mathcal{M}$ parameter space.

Using the cumulative number of templates we split the parameter space into 15 sub-regions, corresponding to contours of equal \mathcal{M} . In figure 5 we show this splitting in $m_1 - m_2$ space.

We populate each sub-region separately in parallel, terminating the Monte Carlo when the number of trials exceeded the square of the current number of templates in the bank. The final number of templates in each sub-region is [495, 346, 352, 375, 404, 408, 416, 432, 462, 479, 489, 517, 517, 532, 510], with mean value 449 and standard deviation 63. The number of templates is approximately 18 times more than predicted by the global ambiguity contours, and as mentioned earlier this is due to the existence of hollow spaces inside these contours. In figure 6 we show one such case. There the global contours overlap but the local maximum of one template falls onto a local minimum of another template.

By filling up each of the sub-regions separately, we over-populate the boundaries between the sub-regions (they were populated from both sides). This overpopulation is not too extreme, $\sim 10\%$, for all 14 boundaries. One can get rid of some templates by computing the overlaps between templates in the vicinity of the boundaries, but this procedure might create ‘holes’: a removed template, a say, which has overlap with template b above MM , might have covered part of the space which is not covered by b . Therefore, one needs to do an additional Monte Carlo to cover the newly created ‘holes’ in the near-boundary regions.

To conclude this section we note that if the efficiency of the method is not a concern, then one can try to populate the whole parameter space without splitting it into subregions or by

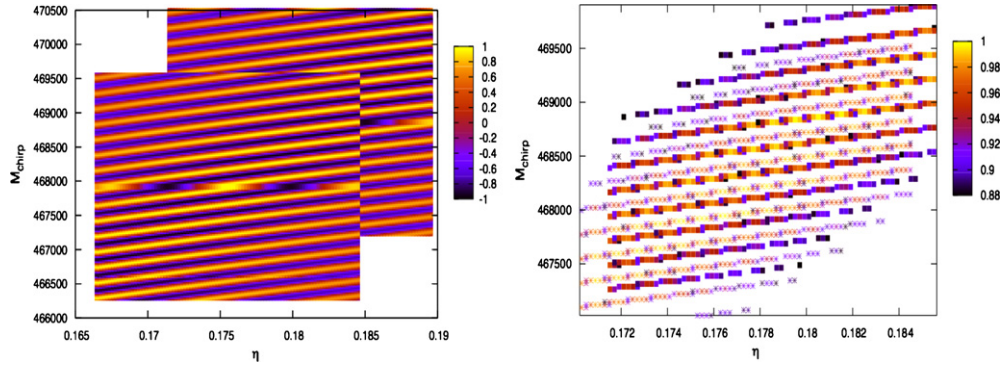


Figure 6. Internal structure of two overlapping global ambiguity contours corresponding to two different templates. Left panel: overlap structure around the central points, the yellow maximum of one template falls onto a local minimum (dark blue) of the other. Right panel: the same as the left panel but here we keep only points with overlap exceeding $MM = 0.88$, crosses and boxes correspond to the ambiguity ellipses of the two different templates $(\{0.175, 467921\}, \{0.18, 468864\})$.

(This figure is in colour only in the electronic version)

making an arbitrary splitting. An alternative would be to populate astrophysically interesting regions of the parameter space (i.e., more likely to contain a signal) with a higher density of templates. However, studying the shapes, sizes and orientations of the global ambiguity contours and local maxima is a very useful exercise that helps us to conduct the search and to understand the results.

3. Application of the stochastic bank to the MLDC

We have filtered the noiseless training data set 1B2.2 [7] through the template bank to test the bank while avoiding bias in the parameter estimation caused by the presence of noise. Even though we have used the training data set, for which the parameters of the signal are known, we conducted a blind search, pretending that we did not know the signal parameters. We did use prior knowledge that there was only one signal with $T_c = 400 \pm 40$ days and SNR between 20 and 100 in a single TDI channel. For the search we made use of the rigid adiabatic approximation to the LISA response (described in appendix A) and used the combined \mathcal{F} -statistic for two channels with uncorrelated noise, X and $Y - Z$ [21].

To search for coalescence time, we use a method similar to the one used in [10]. We start by assuming some T_c , usually at the beginning of the prior, and then compute the correlation for different time lags:

$$C_{h,s}(\tau) = 2 \int \frac{\tilde{s}(f)\tilde{h}^*(f)}{S_n(f)} e^{i2\pi f\tau} df. \quad (3.1)$$

The inner product (1.1) is just a correlation at zero lag. The \mathcal{F} -statistic then becomes an array, $\mathcal{F}(t_k = k\Delta t)$, and we search for its maximum value. In doing this we are trying to fit the intrinsic phase of the signal, assuming the wrong amplitude modulation (as the antenna beam function depends on the choice of t_0 and T_c). Then we use a new value of T_c (the lag at which the maximum of $\mathcal{F}(t_k)$ was found) and repeat the procedure. Usually this converges to a maximum of $\mathcal{F}(t_k)$ at zero lag within a few iterations. This procedure works very well and makes the search for T_c very efficient, but due to the coarseness of the data sampling the

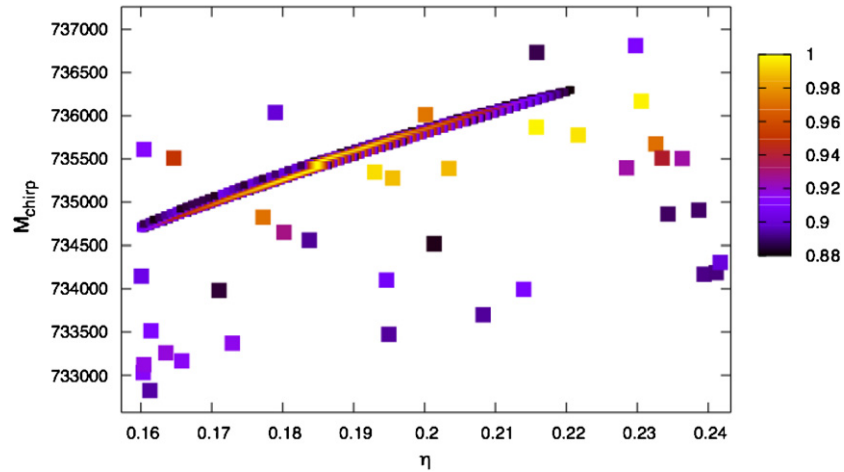


Figure 7. Results of the search: we plot an ambiguity contour at the true point (the overlap value is color-coded) with points of the bank (large boxes) which gave high values of the \mathcal{F} -statistic. For those points color-coding denotes the ratio of the found \mathcal{F} to $\mathcal{F}_{\text{true}}$, the value of the \mathcal{F} -statistic for the template with true parameters.

best estimation of T_c is the nearest multiple of Δt , and for the current search we were using $\Delta t = 120$ s.

Previous studies have shown that the coalescence time correlates with the chirp mass and this (usually undesired) feature helps us to improve the efficiency of the bank. As we have mentioned above and discuss in appendix B ambiguity contours and local maxima depend rather strongly on the choice of T_c for a given MM . However, allowing maximization over T_c (and due to the correlation between T_c and \mathcal{M}) points other than the true point gives a very high value of the \mathcal{F} -statistic. We have tried to illustrate this in figure 7, in which we plot the ambiguity contour (for $MM = 0.88$)³ around the true point and points of the bank which gave a high value of the \mathcal{F} -statistic. Note that those points lie outside the ambiguity ellipse and, at the same time, as we will show soon, those points are off in T_c . We can use the correlation $\mathcal{M} - T_c$ in the second step of the search to find the true parameters of the signal.

This search gives us four parameters: the masses and sky location (with maximization over coalescence time). The best seven points are presented in table 1. The candidates are separated into two groups: the first group corresponds to the true sky location of the signal and the second group shows the antipodal sky position. It is known (see, for example, [10]) that for low-frequency sources ($L\omega_{gw} \ll 1$, where L is the armlength) the LISA response function is quite symmetric with respect to the direct and antipodal sky positions. We have found quite well the sky location and relatively well the chirp mass. We note that the secondary maxima have an \mathcal{F} -statistic value very close to the maximum. So overall our bank did quite well. The candidates can now be passed to the second step of search.

We finish with a comment about the second step of the search. The second step will be similar to that described in [10], but we no longer need frequency annealing. The Metropolis–Hastings stochastic search can be simplified since we are already close to the true point. We will use simulated annealing as it helps us to move the chain off local maxima. Note that we know the global structure of the likelihood (it is the same as the internal structure of the

³ Note that in this case local secondary maxima are lower than 0.88, so the local and global contours are the same.

Table 1. True parameters and the seven best points of the search. The two groups represent the correct and antipodal sky locations. The columns are: primary mass in M_{\odot} , secondary mass in M_{\odot} , symmetric mass ratio, chirp mass in M_{\odot} , ecliptic longitude in radians, ecliptic latitude in radians, time of coalescence in seconds and the value of the \mathcal{F} -statistic

Name	m_1	m_2	η	\mathcal{M}	Long	Lat	T_c	\mathcal{F}
True	1532 832	494 760	0.1845	735 423	5.71	0.158	34 869 697.9	3344.1
Candidate	1265 090	581 679	0.216	735 869	5.70	0.20	34 880 160.0	3338.4
Candidate	1135 330	640 146	0.231	736 169	5.70	0.14	34 882 080.0	3337.7
Candidate	1397 630	534 699	0.200	736 012	5.78	0.03	34 868 520.0	3253.2
Candidate	1214 270	602 610	0.222	735 776	2.60	-0.14	34 882 800.0	3329.4
Candidate	1457 870	515 425	0.193	735 347	2.60	-0.14	34 874 400.0	3325.6
Candidate	1435 940	521 841	0.195	735 278	2.51	-0.14	34 877 160.0	3306.0
Candidate	1368 550	543 364	0.203	735 389	2.51	-0.14	34 879 800.0	3303.1

global ambiguity contours) in that region of space from the bank building stage, so we can use this information to construct an efficient proposal distribution (similar to the ‘island jumping’ described in [10]). So in principle we will need only three proposal distributions: (1) jumps along eigendirections of the Fisher matrix (to search in the vicinity of local maxima), (2) jumps between local maxima, (3) jumps between the direct and antipodal sky locations. We have successfully⁴ implemented this search in the second round of the MLDC and the results are summarized in [22]. Our submission was not the best, other two methods described in the introduction did better. That could be explained by the errors in the implementation mentioned in the footnote. However we have obtained good results for the weakest MBH binary which coalesces outside the observation period. This low frequency signal was weakly affected by the errors and therefore was quite well recovered.

4. Conclusion

In this paper we have presented a new (stochastic) method of constructing a two-dimensional template bank for detecting the gravitational wave signal from inspiralling non-spinning MBH binaries. The main idea is to throw points randomly onto the parameter space and keep only those with overlap below a preset minimal match with all the templates already in the bank. We have shown how to implement such a method efficiently by splitting the parameter space into sub-regions which can be covered with approximately the same number of templates. As a bonus we have also learned about the local and global properties of the likelihood (or overlap) across the parameter space which can be useful in designing search algorithms. Even though we have used this method only in a two-dimensional space, it can be extended to higher dimensional parameter manifolds. We have used this bank to filter a noiseless signal and have shown that the bank is efficient in capturing this signal and making preliminary parameter estimation. Candidates together with information about the properties of the likelihood in the vicinity of those points can be used in the second, follow-up, step of the search.

We should mention that while studying the local and global properties of the overlap function, we have realized the importance of such a study, and have started investigating the global properties of the likelihood surface for signals from Galactic binaries and inspiralling

⁴ To be more precise, this implementation was ‘semi-successful’ as we made an error in reproducing the taper in our model and we used the wrong value of a sidereal year. In addition, due to lack of time, the chains in the second step of search were not long enough.

MBH binaries. There we again use the \mathcal{F} -statistic instead of the overlap. The results of this research will be published in a forthcoming paper.

Acknowledgments

The author would like to thank Bruce Allen, who originally suggested the idea of placing templates stochastically; and to thank Ed Porter for useful discussions. Special thanks to Jonathan Gair for his careful reading of the manuscript and numerous useful suggestions. This work was supported by DLR (Deutsches Zentrum für Luft- und Raumfahrt) and DFG's SFB/TR 7.

Appendix A. Waveform

Here we give a brief summary of the waveform we have used in the search and for constructing the template bank. This model is called 'rigid adiabatic approximation' first suggested in [23], nevertheless we would like to re-derive it here for the consistency with our notations and for the completeness of this paper.

We start with the waveform in the radiative frame, following [24]:

$$h_+^{\text{rad}} = \frac{2M\eta}{D}(M\omega)^{2/3}h_+^o \cos 2(\Phi + \Phi_0) \quad (\text{A.1})$$

$$h_\times^{\text{rad}} = \frac{2M\eta}{D}(M\omega)^{2/3}h_\times^o \sin 2(\Phi + \Phi_0), \quad (\text{A.2})$$

where ω is the instantaneous orbital frequency and Φ is the orbital phase. Note that we are using the restricted PN waveform. The two polarization amplitudes are

$$h_+^o = 1 + \cos^2 i \quad (\text{A.3})$$

$$h_\times^o = -2 \cos i \quad (\text{A.4})$$

and the angle i is the inclination between the orbital angular momentum and the direction to the source in the barycentric frame. It is convenient to use a complex representation of the waveform:

$$h_+ = \frac{M\eta}{D}(M\omega)^{2/3}h_+^o e^{i(\Phi+\Phi_0)} + \text{c.c.} \quad (\text{A.5})$$

$$h_\times = \frac{M\eta}{D}(M\omega)^{2/3}(-ih_\times^o) e^{i(\Phi+\Phi_0)} + \text{c.c.} \quad (\text{A.6})$$

We use a convention for the Fourier transform that agrees with 'fftw':

$$\tilde{h}(f) = \int h(t) e^{-2\pi i f t} dt.$$

In the solar system barycentric frame (SSB)⁵ the waveform is

$$h_+ = h_+^{\text{rad}} \cos 2\psi + h_\times^{\text{rad}} \sin 2\psi \quad (\text{A.7})$$

$$h_\times = -h_+^{\text{rad}} \sin 2\psi + h_\times^{\text{rad}} \cos 2\psi. \quad (\text{A.8})$$

⁵ See for instance [21] for definition of SSB.

We need to compute TDI streams. The single arm response is [25]

$$y_{\text{str}} = (1 + \hat{k}\hat{n}_l)[\Psi_l(t - L - \hat{k}\vec{p}_s) - \Psi_l(t - \hat{k}\vec{p}_r)], \quad (\text{A.9})$$

where

$$\Psi_l = \frac{1}{2} \frac{(\hat{n}_l)_i h^{ij} (\hat{n}_l)_j}{1 - (\hat{k}\hat{n}_l)^2}. \quad (\text{A.10})$$

Here L is the armlength (assumed to be constant), \hat{n}_l is a unit vector along the interferometer's arm, $\hat{k} = -\hat{n}$ is the direction of propagation of the GW, and $\vec{p}_j = \vec{R} + \vec{q}_j$ is the position of the j th spacecraft in the SSB, \vec{R} is the position of the guiding center and \vec{q}_j is the spacecraft's position with respect to the guiding center. The field of the gravitational wave is

$$h^{ij} = h_+ \mathcal{E}_+^{ij} + h_\times \mathcal{E}_\times^{ij} \quad (\text{A.11})$$

$$\mathcal{E}_+^{ij} = u^i u^j - v^i v^j, \quad \mathcal{E}_\times^{ij} = u^i v^j + u^j v^i. \quad (\text{A.12})$$

Here \hat{u} , \hat{v} define the polarization basis in the SSB:

$$-\hat{k} = \hat{n} = \{\sin \theta_S \cos \phi_S, \sin \theta_S \sin \phi_S, \cos \theta_S\} \quad (\text{A.13})$$

$$\hat{u} = \{\cos \theta_S \cos \phi_S, \cos \theta_S \sin \phi_S, -\sin \theta_S\} \quad (\text{A.14})$$

$$\hat{v} = \{\sin \phi_S, -\cos \phi_S, 0\}. \quad (\text{A.15})$$

Taking the above into account we get

$$\frac{1}{2} (\hat{n}_l)_i h^{ij} (\hat{n}_l)_j = U_l h_+ + V_l h_\times \quad (\text{A.16})$$

$$= \frac{M\eta}{D} (M\omega)^{2/3} e^{2(\Phi+\Phi_0)} [U_l F_u + V_l F_v] + \text{c.c.} \quad (\text{A.17})$$

and

$$U_l = \frac{1}{2} [(\hat{n}_l \hat{u})^2 - (\hat{n}_l \hat{v})^2], \quad V_l = (\hat{n}_l \hat{u})(\hat{n}_l \hat{v}) \quad (\text{A.18})$$

$$F_u = h_+^o \cos 2\psi - i h_\times^o \sin 2\psi, \quad F_v = -h_+^o \sin 2\psi - i h_\times^o \cos 2\psi. \quad (\text{A.19})$$

Then one can compute the response for a particular TDI combination using the phase expansion around the SSB time which leads to the following expression for a single link:

$$y_{\text{str}} = i \frac{M\eta}{D} (M\omega)^{2/3} [U_l F_u + V_l F_v] e^{i2[\Phi(t)+\Phi_0-\omega\hat{k}\vec{R}]} (\omega L) \times \text{sinc}[\omega L(1 - \hat{k}\hat{n}_l)] e^{-i\omega(L+\hat{k}(\vec{p}_s+\vec{p}_r))} + \text{c.c.} \quad (\text{A.20})$$

This approximation is called the rigid adiabatic approximation. In particular for the X -channel (unequal arm Michelson) we have

$$X^{gw} = \frac{M\eta}{D} (M\omega)^{2/3} e^{i2[\Phi(t)+\Phi_0-\omega\hat{k}\vec{R}]} \sum_{j=2,3} [U_j F_u + V_j F_v] \chi_j^X \quad (\text{A.21})$$

actually, only χ_j is different for different TDIs. For X we can write it explicitly:

$$\chi_2^X = -2x \sin x \left\{ \text{sinc} \left[\frac{x}{2} (1 - \hat{k}\hat{n}_1) \right] e^{-ix} + \text{sinc} \left[\frac{x}{2} (1 + \hat{k}\hat{n}_1) \right] \right\} e^{-i\frac{x}{2} [3+\hat{k}(\vec{q}_0+\vec{q}_2)]} \quad (\text{A.22})$$

$$\chi_3^X = 2x \sin x \left\{ \text{sinc} \left[\frac{x}{2} (1 - \hat{k}\hat{n}_2) \right] + \text{sinc} \left[\frac{x}{2} (1 + \hat{k}\hat{n}_2) \right] e^{-ix} \right\} e^{-i\frac{x}{2} [3+\hat{k}(\vec{q}_1+\vec{q}_0)]}.$$

In the long wavelength limit $x = L\omega_{gw} \ll 1$, those become

$$\chi_2^X = -4x^2, \quad \chi_3^X = 4x^2.$$

We have used the long wavelength expression for constructing the bank, and we have used the expressions (A.22) for computing the \mathcal{F} -statistic.

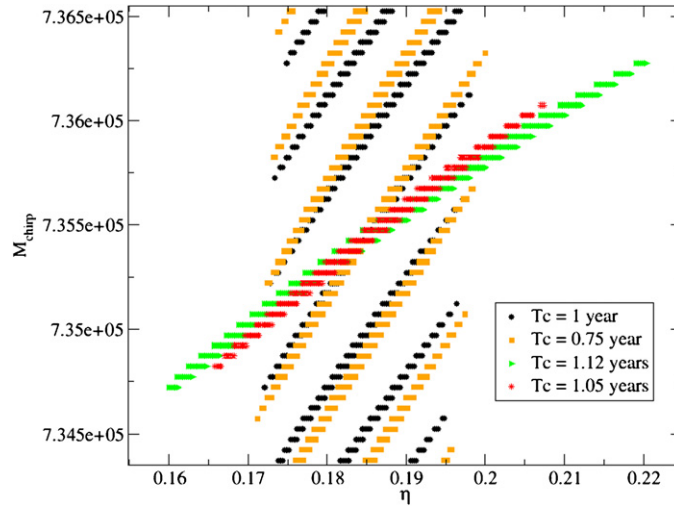


Figure B1. Ambiguity ellipses for different coalescence times.

Appendix B. Effect of coalescence time on the ambiguity contours

We have computed ambiguity contours around the true point for different fixed times of coalescence. In other words, we kept the sky location and extrinsic parameters constant and plotted the ambiguity contour in the η - \mathcal{M} plane for different values of T_c . The results are presented in figure B1. One can notice several important points: (i) ellipses of local maxima increase in size as we go to larger T_c , (ii) ellipses of local maxima rotate and the distance between them changes, (iii) the maximum overlap for the secondary maxima decreases as we increase T_c , (iv) this is not obvious from figure B1, but the global contour also slightly rotates.

All these features can be easily explained: by changing T_c and keeping a fixed duration of the observation we effectively change the frequency range of the signal: by increasing T_c we are observing the low-frequency part of the inspiral only. We know that the accuracy in parameter estimation drops for a low-frequency source; therefore we see an increase in the size of the local maxima for larger T_c . Low frequencies also correspond to the less relativistic regime of the inspiral, and the phase of the wave is governed largely by the leading order term which depends on \mathcal{M} only. Therefore, for large T_c , the uncertainty in η increases and the ellipses are more inclined toward the η -axis. Once we start to see the end of inspiral (which is ≥ 1 mHz), the majority of the SNR comes from the last ten-to-hundred cycles (the signal reaches the most sensitive bandwidth of LISA, and the intrinsic amplitude of the signal increases rapidly). Similarly for the overlap, the main contribution comes from the end of the inspiral. If we observe the whole inspiral, the last \sim hundred cycles can be fit by many templates with different parameters, therefore we see a set of secondary maxima with a slowly damping strength (the maximal overlap within the secondary ellipses).

This figure also suggests that for $MM = 0.88$ the conservative choice (smallest ambiguity contour) corresponds to $T_c \approx 1.05$ years. But this choice would need a significantly larger number of templates, which we believe is not necessary. This is because, as we discussed in the main body of the paper, of the maximization over T_c and the correlation of T_c with \mathcal{M} . What is more important is to take into account this correlation during the second step of the search to identify the true parameters (primary maximum).

References

- [1] Zel'dovich Y and Novikov I 1964 *Dokl. Akad. Nauk. SSSR* **158** 811
- [2] Tremaine S *et al* 2002 *Astrophys. J.* **574** 740 (Preprint [astro-ph/0203468](#))
- [3] Hopkins P F *et al* 2006 *Astrophys. J. Suppl.* **163** 1 (Preprint [astro-ph/0506398](#))
- [4] Rodriguez C *et al* 2006 *Astrophys. J.* **646** 49 (Preprint [astro-ph/0604042](#))
- [5] Danzmann K *et al* 1998 *LISA Laser Interferometer Space Antenna, Pre-Phase A Report* (Max-Planck-Institute fur Quantenoptic, Report MPQ 233)
- [6] Blanchet L, Iyer B R, Will C M and Wiseman A G 1996 *Class. Quantum Grav.* **13** 575 (Preprint [gr-qc/9602024](#))
- [7] Arnaud K A *et al* 2006 *AIP Conf. Proc.* **873** 619 (Preprint [gr-qc/0609105](#))
- [8] Arnaud K A *et al* (Mock LISA Data Challenge Task Force) 2006 *AIP Conf. Proc.* **873** 625 (Preprint [gr-qc/0609106](#))
- [9] Cornish N J and Porter E K 2007 *Class. Quantum Grav.* **24** S501 (Preprint [gr-qc/0701167](#))
- [10] Cornish N J and Porter E K 2007 *Class. Quantum Grav.* **24** 5729 (Preprint [gr-qc/0612091](#))
- [11] Cornish N J and Porter E K 2007 *Phys. Rev. D* **75** 021301 (Preprint [gr-qc/0605135](#))
- [12] Brown D A, Crowder J, Cutler C, Mandel I and Vallisneri M 2007 *Class. Quantum Grav.* **24** S595 (Preprint [arXiv:0704.2447](#))
- [13] Babak S, Balasubramanian R, Churches D, Cokelaer T and Sathyaprakash B S 2006 *Class. Quantum Grav.* **23** 5477 (Preprint [gr-qc/0604037](#))
- [14] Prix R 2007 *Class. Quantum Grav.* **24** S481 (Preprint [arXiv:0707.0428](#))
- [15] Harry I W, Fairhurst S and Sathyaprakash B S 2008 *Class. Quantum Grav.* **25** 184027 (Preprint [arXiv:0804.3274](#))
- [16] Messenger C, Prix R and Papa M A 2008 in preparation
- [17] Arnaud K A *et al* 2007 *Class. Quantum Grav.* **24** S551 (Preprint [gr-qc/0701170](#))
- [18] Jaranowski P, Krolak A and Schutz B F 1998 *Phys. Rev. D* **58** 063001 (Preprint [gr-qc/9804014](#))
- [19] Balasubramanian R, Sathyaprakash B S and Dhurandhar S V 1996 *Phys. Rev. D* **53** 3033 (Preprint [gr-qc/9508011](#))
- [20] Owen B J and Sathyaprakash B S 1999 *Phys. Rev. D* **60** 022002 (Preprint [gr-qc/9808076](#))
- [21] Krolak A, Tinto M and Vallisneri M 2004 *Phys. Rev. D* **70** 022003 (Preprint [gr-qc/0401108](#))
- [22] Babak S *et al* 2008 *Class. Quantum Grav.* **25** 114037 (Preprint [arXiv:0711.2667](#))
- [23] Rubbo L J, Cornish N J and Poujade O 2004 *Phys. Rev. D* **69** 082003 (Preprint [gr-qc/0311069](#))
- [24] Finn L S and Chernoff D F 1993 *Phys. Rev. D* **47** 2198 (Preprint [gr-qc/9301003](#))
- [25] Armstrong J W, Estabrook F B and Tinto M 1999 *Astrophys. J.* **527** 814