A constrained Metropolis–Hastings search for EMRIs in the Mock LISA Data Challenge 1B

Jonathan R Gair[1]**, Edward Porter**[2]**, Stanislav Babak**[2] **and Leor Barack**[3]

[1] Institute of Astronomy, University of Cambridge, Cambridge CB3 0HA, UK
[2] Albert Einstein Institute, Am Muehlenberg 1, D-14476 Golm, Germany
[3] School of Mathematics, University of Southampton, Southampton SO17 1BJ, UK

E-mail: jgair@ast.cam.ac.uk

Abstract

We describe a search for the extreme-mass-ratio inspiral sources in the Round 1B Mock LISA Data Challenge data sets. The search algorithm is a Monte Carlo search based on the Metropolis–Hastings algorithm, but also incorporates simulated, thermostated and time annealing, plus a harmonic identification stage designed to reduce the chance of the chain locking onto secondary maxima. In this paper, we focus on describing the algorithm that we have been developing. We give the results of the search of the Round 1B data, although parameter recovery has improved since that deadline. Finally, we describe several modifications to the search pipeline that we are currently investigating for incorporation in future searches.

PACS numbers: 04.25.Nx, 04.30.Db, 04.80.Cc, 04.80.Nn, 95.55.Ym, 95.85.Sz

1. Introduction

The inspirals of stellar mass compact objects into supermassive black holes in the centres of galaxies—extreme-mass-ratio inspirals (EMRIs)—are one of the most exciting potential sources of gravitational waves for the planned Laser Interferometer Space Antenna (LISA). The detection of such sources in the LISA data stream and parameter estimation for them is a very challenging technical problem, however, as the instantaneous amplitude of a typical signal is an order of magnitude below the noise fluctuations in the detector. Moreover, the long duration of the signals (LISA will detect up to 10^5 waveform cycles in an observation) and the large parameter space of possible sources (an EMRI signal depends on 14 parameters) makes fully coherent matched filtering computationally impossible [1].

Several possible algorithms have been considered for EMRI detection. In order to compute initial event rate estimates, a semi-coherent algorithm was proposed, in which the data stream would be divided up into short (2 or 3 weeks) segments, that would be searched coherently via

matched filtering. The signal-to-noise ratio (SNR) would then be built up in a second stage via incoherent summation of power along trajectories through the coherent segments [1]. This algorithm was designed to make full use of available computing power and its effectiveness has not yet been demonstrated practically. Time–frequency techniques have also been explored [2–6] and have been shown to be able to both detect and recover parameters [6] when used to search for single, high-SNR EMRIs in instrumental noise. Their effectiveness is likely to be significantly reduced when confronted with more realistic situations, in which there are multiple sources overlapping in time and frequency. However, these algorithms may provide a useful first step in a hierarchical search.

Markov chain Monte Carlo (MCMC) methods may provide a way to perform matched filtering more efficiently, without requiring a large template bank of possible signals. We follow the convention in the literature and call our search a Metropolis–Hastings Monte Carlo (MHMC) search rather than an MCMC search since it is not actually Markovian (see section 2.2). Both MCMC and MHMC methods have been explored by various groups in the context of LISA data analysis and have been shown to be very effective when searching for toy models [7, 8], for white-dwarf binaries [9, 10] and for single or multiple supermassive black hole binaries [11–15]. The use of an MCMC technique for EMRI searches was explored by Stroeer *et al* [16], based on a highly simplified model of the EMRI waveform. In the context of the Mock LISA Data Challenges (MLDCs) [17], Monte Carlo methods have been used by ourselves and one other group [18] to search for the EMRI sources in the Round 2 [19] and Round 1B [20] data sets. In this paper, we describe the algorithm that we have developed for the EMRI searches, and the performance of the algorithm on the Round 1B challenge data sets. Our search code was adapted from the search code developed by Cornish and Porter [11, 13–15] for supermassive black hole binary searches, but we have incorporated a significant number of refinements that are specific to the EMRI problem.

The paper is organized as follows. In section 2 we describe the search algorithm, including a description of the waveform model, the Metropolis–Hastings search engine and various refinements we have tried for the EMRI problem. In section 3 we present the results that we had at the time of the MLDC Round 1B deadline (December 2007) and compare these to the true source parameters. Finally, in section 4 we discuss planned future refinements of the search algorithm.

2. Search algorithm

2.1. The waveform model

The EMRI sources in the MLDC releases to date were constructed using the *analytic kludge* (AK) model of Barack and Cutler [21]. In this model, the gravitational waveforms describe emission from a Keplerian orbit, but with perihelion and orbital-plane precessions imposed by precessing the observer about the source with rates taken from post-Newtonian expressions. The orbital parameters are also evolved over time using post-Newtonian prescriptions to account for radiation reaction. The EMRI waveforms have emission at multiple frequencies, corresponding to harmonics of the fundamental frequencies of the orbit—harmonics of the orbital frequency, ν, arise from the eccentricity of the orbit, e; harmonics of the perihelion precession rate, $\dot{\gamma}/2\pi$, arise from this precession; and harmonics of the orbital-plane precession rate, $\dot{\alpha}/2\pi$, are present due to the inclination, λ, of the orbital plane relative to the equatorial plane of the black hole. The frequency of a given waveform harmonic is given by three integers, (n, l, k), as $f = n\nu + l\dot{\gamma}/2\pi + k\dot{\alpha}/2\pi$. The AK waveforms are purely quadrupole in nature, and so $|l|, |k| \leqslant 2$, but n is unrestricted. However, the eccentricity at plunge of

Table 1. Percentage difference between the parameters of the best-fit truncated waveform and the parameters of the true AK waveform. The true AK waveform is the 1.3.2 training waveform from MLDC Round 2, for which the parameters are $\ln(m/M_\odot) = 2.3338$, $\ln(M/M_\odot) = 15.421$, $\ln(\nu_0/\text{Hz}) = -7.9399$, $D_L/\text{Gpc} = 0.805\,33$, $e_0 = 0.187\,65$, $\chi = 0.684\,97$, $\cos(\lambda) = -0.439\,60$, $\cos(\beta) = 0.967\,37$, $\phi_S = 3.6238$, $\cos(\theta_K) = -0.554\,34$, $\phi_K = 4.3216$, $\Phi_0 = 3.3913$, $\gamma_0 = 6.1502$ and $\alpha_0 = 3.2400$.

Parameter	$\ln(m/M_\odot)$	$\ln(M/M_\odot)$	$\ln(\nu_0/\text{Hz})$	D_L/Gpc	e_0	χ	$\cos(\lambda)$
Error (%)	0.006	0.0008	0.000 04	7	0.07	0.006	0.06
Parameter	$\cos(\beta)$	ϕ_S	$\cos(\theta_K)$	ϕ_K	Φ_0	γ_0	α_0
Error (%)	0.2	0.06	0.06	0.3	3	3	0.5

the MLDC sources is limited to $[0.15, 0.25]$, and so harmonics with $n \geqslant 6$ are generally weak.

In our search code, to speed up waveform evaluation, we use a truncated version of the AK model. We include only the $n \leqslant 5$ and $l = 2$ harmonics in the waveform model (harmonics with $l \neq 2$ are significantly suppressed relative to the $l = 2$ harmonics). In addition, we expand the Bessel functions that appear in the model [21] in powers of eccentricity, and keep only the three leading terms in the expansion of $J_n(ne)$ for each n. We include the full LISA TDI response function to account for detector motions, and use parameters evaluated at plunge to characterize the waveforms (this is in contrast to the MLDC convention, which is to specify parameters at the start of the observation). The resulting waveforms are faithful approximations to the full AK waveforms. The overlap between an AK and truncated waveform, evaluated at the same waveform parameters, is at worst 90%, and is typically 95%. The overlap tends to be higher for sources with higher mass central black holes, for which the emission is mostly at lower frequencies (for $M \sim 10^7 M_\odot$ the overlap always exceeds 96%). The template parameter error, i.e., the difference between the parameters of the best-fit truncated waveform and the parameters of the true AK waveform, is also relatively small. In table 1 we list the parameter errors for one of the MLDC Round 2 training sources. These results were obtained by starting an MCMC chain at the true parameter values and allowing it to evolve to evaluate the posterior. The difference between the mean of the recovered posterior and the true parameter values provided an estimate of the bias in our truncated model. We note that since the data stream we were searching included noise we expected and saw a noise-induced bias in the parameter estimation. However, we could distinguish this noise bias from the model errors. The parameter offset values in table 1 are typical for MLDC-type sources. The parameters are the same as those used for the MLDC—compact object mass (m), central black hole mass (M), initial orbital frequency (ν_0), luminosity distance (D_L), initial eccentricity (e_0), central black hole spin (χ), ecliptic latitude and longitude (β, ϕ_S), orientation of central black hole spin (θ_K, ϕ_K), orbital inclination (λ) and three initial orbital phases ($\Phi_0, \gamma_0, \alpha_0$). For most parameters, the error is at most 2σ, where σ is the noise-induced uncertainty in the parameter, as estimated from the width of the posterior. The error is somewhat larger for the initial phase angles, as the effect of the truncation accumulates over the observation, but these parameters are uninteresting astrophysically. Overall, the truncated model provides an estimate of the parameters that is sufficiently close to the true parameters to ensure that a follow-up refinement with full AK waveforms would be quick.

2.2. Metropolis–Hastings Monte Carlo algorithm

Our search engine is based in the Metropolis–Hastings algorithm, which works as follows: given a data set $s(t)$ and a set of templates $h(t; \vec{x})$, we randomly choose a starting point, $\vec{x}$,

in the parameter space. We then propose a jump to another point, $\vec{y}$, in the space by drawing from a certain proposal distribution, $q(\vec{y}|\vec{x})$ (see below), and evaluate the Metropolis–Hastings ratio

$$H = \frac{\pi(\vec{y})p(s|\vec{y})q(\vec{x}|\vec{y})}{\pi(\vec{x})p(s|\vec{x})q(\vec{y}|\vec{x})}. \tag{1}$$

Here $\pi(\vec{x})$ are the priors of the parameters, which, in our analysis, were taken to be uniform distributions within the ranges allowed by the MLDC. The function $p(s|\vec{x})$ is the likelihood

$$p(s|\vec{x}) = C\, e^{-\langle s - h(\vec{x})|s - h(\vec{x})\rangle/2}, \tag{2}$$

where C is a normalization constant. This jump is then accepted with probability $\alpha = \min(1, H)$, otherwise the chain stays at $\vec{x}$.

If the proposal distribution was independent of the step number, the algorithm would be Markovian. However, the MHMC algorithm we employ is not, since we use a variety of purposely directed proposal distributions that allow a range of jumps of different size and type in the parameter space. We also implement several annealing schemes to encourage movement of the chain and use time sliding to search automatically over the plunge time (which we use as a parameter instead of the initial frequency, ν_0). We are using the MHMC algorithm as a search code to find the unknown parameters of the signal. Making the search non-Markovian allows more rapid convergence to the source parameters, at the expense of no longer being able to construct the posterior from the chain state distribution. In a future implementation of the pipeline, we will include a final Markov Chain stage to recover the posterior once the source parameters have been approximately identified. This was not implemented for the searches described here.

2.3. Proposal distributions

Our primary proposal distribution is a multi-variate Gaussian, which is constructed as a product of Gaussian distributions in each eigendirection of the Fisher information matrix (FIM), Γ_{ij}. The distribution in each eigendirection is taken to have zero mean and a standard deviation of $\sigma_i = 1/\sqrt{DE_i}$. Here D is the dimensionality of the search space (in this case, $D = 13$, as we use normalized search templates) and E_i is the corresponding eigenvalue of $\Gamma_{\mu\nu}$. The FIM is computed numerically, using the same truncated AK model that we employ as the search template, but with the additional simplification that the detector response is modelled using the low-frequency approximation (as in the original AK paper [21]), rather than via the full TDI response.

While the FIM-based proposal is used at most steps, we also periodically draw a proposed point from one of several other proposal distributions:

- *Scaled uniform jumps*—the proposal distribution is taken to be uniform within the priors for each parameter. This forces the chain to explore other, widely separated, regions of the parameter space.
- *Scaled FIM jumps*—this proposal is based on the FIM as for the standard proposal, but the proposed jumps are artificially reduced in size by a factor of 10. These proposals are almost always accepted, forcing the chain to move slightly away from secondaries and hence encouraging movement.
- *Antipodal sky position*—at low frequencies, the sky positions (β, ϕ_S) and $(-\beta, \phi_S \pm \pi)$ are equivalent in terms of the detector response. In searches for black hole binaries, it was found that the chains could often become locked on the wrong sky position, so we include a proposal that moves the chain to the antipodal sky position to avoid this problem.

- *Intrinsic/extrinsic/phase only jumps*—the waveforms can be divided into *intrinsic* parameters $(m, M, \nu_0, e_0, \chi, \lambda)$, *extrinsic* parameters $(D_L, \beta, \phi_S, \theta_K, \phi_K)$ and *phase offsets* $(\Phi_0, \gamma_0, \alpha_0)$. Intrinsic parameters affect the frequency and phase evolution of the different harmonics, while extrinsic parameters only affect how this is projected into a detector response, and the phase offsets only define the relative phase of the different harmonics at one fiducial time. Proposing jumps in extrinsic/phase offset parameters only was designed to improve the fit while keeping the harmonic frequencies fixed (see section 2.5).

We note that no matter which proposal we use, all the waveform parameters are updated, i.e., we do not use Gibbs sampling to update one parameter at a time. Typically, we draw from each of the alternative proposals every few tens of chain points, but these sampling frequencies are tunable parameters which we have not yet optimized for the EMRI search, and so we do not quote specific numbers here.

2.4. Simulated, thermostated and time annealing

As the likelihood surface for EMRIs is full of secondary maxima, it is very easy for the search chains to get stuck. To ensure acceptable movement in the chain, we use simulated annealing to heat the likelihood surface [13, 15]. This smooths and flattens any bumps on the likelihood surface, which ensures the chain has a greater chance of escaping secondaries and moving uphill towards the primary peak. In the definition of the likelihood, equation (2), there is a factor of 2 in the denominator of the exponent. We replace this factor with a 'heat' parameter Θ, defined by

$$\Theta = 2 \times \begin{cases} 10^{\xi(1-\frac{i}{T_c})} & 0 \leqslant i \leqslant T_c \\ 1 & i > T_c, \end{cases} \tag{3}$$

where $\xi > 0$ is the heat-index defining the initial heat, i is the step number of the chain and T_c is the cooling schedule (i.e., the number of chain steps over which the cooling takes place). As the success of the choice of initial heat is only known *a posteriori*, we are still investigating the optimal heating scheme.

To mitigate potential problems, we also employed a thermostated heating scheme as used in [15]. We define

$$\delta = 2 \times \begin{cases} 1 & 0 \leqslant \text{SNR} \leqslant \rho_c \\ \left(\frac{\text{SNR}}{\rho_c}\right)^2 & \text{SNR} > \rho_c, \end{cases} \tag{4}$$

and take the heat to be the maximum of δ and Θ defined by equation (3). The thermostated heating scheme pumps heat into the system once we reach $\text{SNR} = \rho_c$. This should make the chain move uphill faster, once we have begun to match the signal. This scheme is run in conjunction with the simulated annealing. While we continue to calculate both heats, we always use the heat that is highest in the search. For this study we chose $\rho_c = 15$.

The final annealing scheme we employ is time annealing, which is similar to the frequency annealing described in [15]. The cost of evaluating the likelihood and the FIM for the proposal distribution depends on the length of the waveform template being used. It is inefficient to use full length (2 years) templates initially, when the parameter values are poorly constrained. We therefore start off with shorter templates and increase the length of the template, t_{dur}, as the chain progresses

$$t_{\text{dur}} = \begin{cases} t_{\max}(t_{\min}/t_{\max})^{(1-\frac{i}{N_f})} & t < t_{\max} \\ t_{\max} & t \geqslant t_{\max}, \end{cases} \tag{5}$$

where i is the number of steps in the chain, N_f the total number of iterations in the time annealing chain, t_{min} is the minimum length of template and t_{max} is the maximum length of the template. The aim of the time annealing is to use shorter templates while the chain is taking big jumps exploring the parameter space, and then to increase the template length and refine the parameter determination once the chain has settled in the vicinity of the true parameters. For this analysis, we used $t_{min} = 4$ months to ensure a reasonable SNR ($\geqslant 20$) in the shortest templates employed, and we took $N_f = 10\,000$ and $t_{max} = 1$ year, half the length of the MLDC data release.

In our implementation of the search, we started by running a 5000 iteration chain with t_{dur} fixed at 4 months, but using simulated annealing and thermostated heating. After 5000 iterations we began a 10 000 iteration time-annealing scheme, using thermostated heating in conjunction but not simulated annealing. Once the time-annealing phase was complete, we cooled the surface down over 5000 iterations during a final simulated annealing phase according to the scheme detailed by equation (3) without thermostated heating. We then allowed the chain to run as a standard MCMC with unit heat for 80 000 iterations to obtain a final refinement of the parameters.

2.5. Harmonic identification

The EMRI likelihood space is very complicated and is characterized by many secondary maxima. It is consequently very easy for the chain to get stuck on a secondary. This was a major concern in our analysis of the MLDC Round 2 data, and again in the initial analysis of the Round 1B data. A chain started at a random point would lock very quickly onto a secondary of the signal, but would then not move far away from that point. However, chains seeded at different initial points would lock onto different secondaries. A secondary typically has one or two harmonics that overlap with harmonics of the true signal for a section of the data stream, and we can use that information to determine the true parameters. Given the parameters of a secondary, one can construct the cumulative overlap (in the frequency domain) of each harmonic of the secondary with the data stream. If the overlap is significant, the frequency range in which the overlap accumulates indicates the section of that secondary harmonic which matches a harmonic of the true signal. This is illustrated in the top panel of figure 1.

In order to exploit this information, as a preliminary stage of the search we ran several ($\sim$20) chains, seeded at different points in parameter space, stopped the chains after $\sim$1000 steps and then analysed the harmonic content of the highest SNR point found in each chain. Different chains would typically find different harmonics, and so combining this information allows a picture of the signal structure to be constructed. This is illustrated in the bottom panel of figure 1. This information can be used in the search in several ways—as a veto of secondaries (any parameter space point that does not match these harmonics cannot be the primary maximum); to fold into parameter priors (reject proposed steps that move harmonics too far away from these tracks); or as a constraint when proposing steps for the chain.

2.6. Constrained jumps

It is possible to modify the proposal distribution to ensure certain constraints are satisfied, e.g., the frequency of a given harmonic at a given time. Given a set of constraints, $\{f_i(\vec{x}) = 0\}$ for $i = 1, \ldots, N$, we can define a set of unit vectors orthogonal to the constraint surfaces, $\vec{e}_i = \nabla f_i / |\nabla f_i|$, and decompose a step in parameter space, $\delta \vec{x}$, into a piece that maintains the constraints, $\delta \vec{\tilde{x}}$, and a piece in the space spanned by the $\vec{e}_i$, $\delta \vec{x} = \delta \vec{\tilde{x}} + a_i \vec{e}_i$. The template
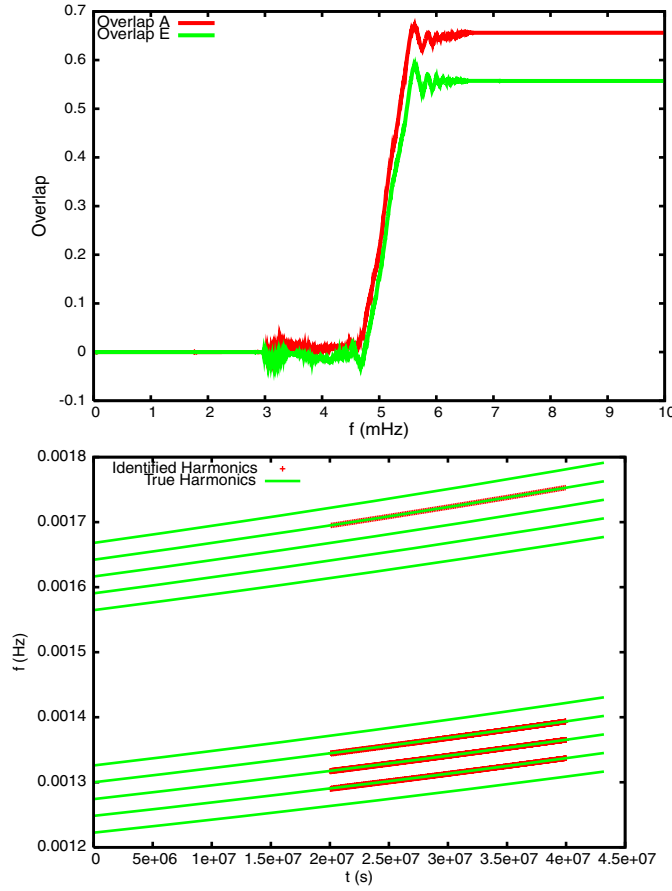
Figure 1. *Top panel*: cumulative overlap of a given harmonic with the data stream, as a function of frequency. The curves show the overlap of the A and E TDI data streams, respectively. The overlap accumulates when the harmonic frequency is $4.5\,\mathrm{mHz} \lesssim f \lesssim 5.5\,\mathrm{mHz}$, which indicates the time range where it matches a harmonic of the true signal. *Bottom panel*: true harmonics (lines) and identified harmonics (crosses) for the MLDC Round 1B training source 1B.3.2.

mismatch, $\mathcal{M}$, at the new point for a jump that maintains the constraints is therefore given in terms of the FIM, Γ_{ij}, by

$$
\begin{aligned}
2\mathcal{M} &= \Gamma_{ij}\delta\tilde{x}^i\delta\tilde{x}^j = \tilde{\Gamma}_{ij}\delta x^i\delta x^j \\
&= \left(\Gamma_{ij} - 2\Gamma_{mj}C_{lk}e_k^i e_l^m + \Gamma_{mn}C_{lk}C_{pq}e_k^i e_q^j e_l^m e_p^n\right)\delta x^i\delta x^j
\end{aligned}
\tag{6}
$$

in which superscripts indicate vector components, $C_{ij} = (A^{-1})_{ij}$ for $A_{ij} = \vec{e}_i \cdot \vec{e}_j$ and repeated indices are summed as usual. The matrix $\tilde{\Gamma}_{ij}$ can be used in place of Γ_{ij} to generate proposed steps that maintain the constraints. The matrix $\tilde{\Gamma}_{ij}$ has N zero eigenvalues with corresponding eigenvectors $\{\vec{e}_i\}$, which are ignored when constructing the proposed step.

We implemented constrained jumps in the search, using the results of the harmonic analysis described in section 2.5 to determine the frequencies of several harmonics and their rates of change at a specified time. In practice, we found it best to use the frequencies of three harmonics, and the time derivative of the dominant harmonic as the constraints. Additional frequencies or derivatives did not provide any extra information. For the AK model, it is
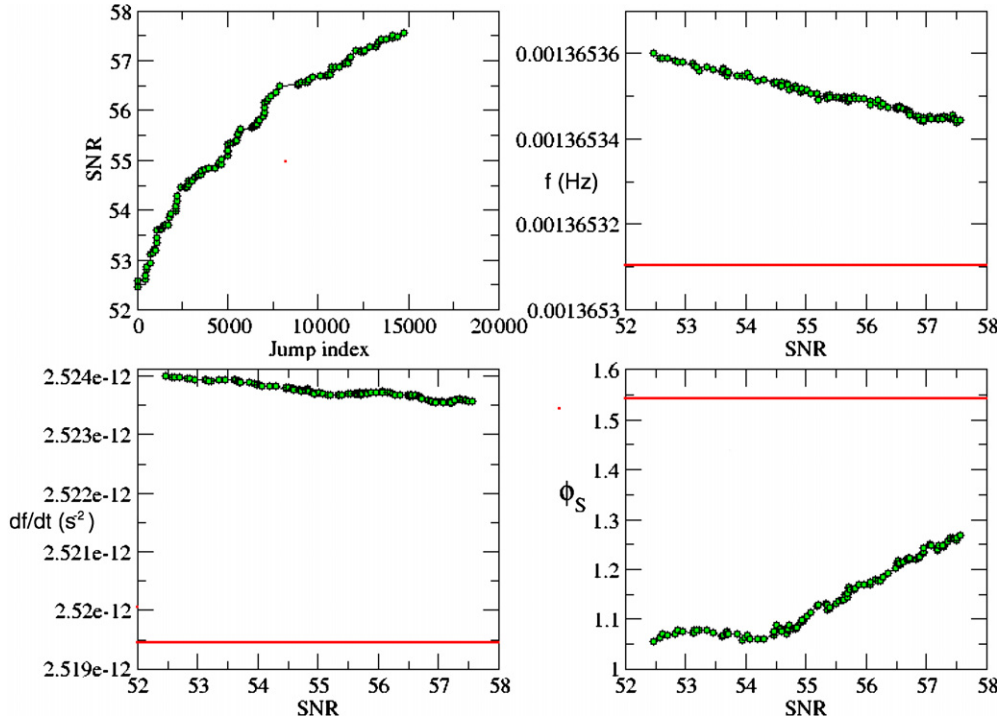
Figure 2. SNR as a function of chain index (top-left); frequency of dominant harmonic versus SNR (top-right); rate of change of frequency of dominant harmonic versus SNR (bottom-left); and sky longitude versus SNR (bottom-right), in a partially constrained search for the MLDC Round 1B.3.2 challenge source. The chain is moving steadily towards the true parameter values, indicated by horizontal lines. This run did not employ either simulated annealing or time-annealing.

easy to translate such constraints into parameter values and hence reduce the six-dimensional intrinsic parameter space—M, m, S, λ, e_0, ν_0—to a two-dimensional space spanned by e_0 and λ. We actually used this dimensional reduction in our implementation of the search, although the general expression (6) will be necessary for a generic waveform model.

The constrained search was able to rapidly improve the SNR from the initial guess and if the constraints were specified exactly (using the MLDC training data), the chain moved steadily to the correct point. However, the harmonic analysis was not able in general to determine precise values for the frequencies. An error of $\sim 10^{-7}$ Hz in one frequency leads to a dephasing after ~ 4 months, which limits the usefulness of the fully constrained search. We found it more effective to use a partially constrained search, i.e., using the harmonic analysis to estimate frequencies and frequency derivatives at a certain time *with estimated uncertainties*. It is possible to reparametrize the waveform in terms of the values of the three fundamental frequencies at a certain time, plus the time derivative of the dominant harmonic at the same time. An MHMC chain can then be run on this alternative parameter space, using tight priors on those frequencies assigned using the frequency uncertainty estimates. This can also be done directly in the physical parameter space by using the constrained FIM, $\tilde{\Gamma}_{ij}$, and additionally allowing small jumps in the directions $\vec{e}_i$ orthogonal to the constraint surface. The results of one partially constrained search are illustrated in figure 2. This type of search was found to be much more effective, although our initial implementation of the proposal was inefficient which led to a low acceptance rate.

Table 2. Comparison of best recovered parameters at time of MLDC submission to true parameters.

	Source	S/M^2	$m/M_\odot$	$M/(10^6 \, M_\odot)$	ν_0(mHz)	e_0	λ
1B.3.1	True	0.698 16	10.296	9.5180	0.192 04	0.214 38	0.439 46
	Recovered	0.636 24	10.500	10.359	0.187 11	0.158 10	0.508 00
1B.3.2	True	0.637 96	9.7711	5.2156	0.342 28	0.207 91	1.4358
	Recovered	0.639 71	9.7751	5.2076	0.342 23	0.209 41	1.4399
1B.3.3	True	0.533 26	9.6973	5.2197	0.342 57	0.199 27	0.928 22
	Recovered	0.596 55	10.193	5.2344	0.342 36	0.196 47	0.758 82
1B.3.4	True	0.625 14	10.105	0.955 80	0.851 44	0.450 58	1.6707
	Recovered	0.631 04	10.085	1.0439	0.795 10	0.440 77	2.1837
1B.3.5	True	0.658 30	9.7905	1.0334	0.832 18	0.426 91	2.3196
	Recovered	0.677 01	9.8849	0.978 72	0.833 90	0.429 50	2.5092

3. Round 1B results

At the time of the MLDC submission deadline (December 2007), we were still developing the details of the search outlined above, and therefore just submitted the maximum *a posteriori* (MAP) parameters that had been found in the search by that time. We used the MAP values as we were sure at this stage that we were either still moving towards the true solution or stuck on a secondary. It therefore made no sense to use the mean of a posterior to determine the best-fit parameters. Our strategy involved starting 10 chains at different random start points for each challenge. Due to time constraints we needed to stop the MCMC chains after between 35 000 and 40 000 iterations out of 80 000. We used the MAP values from each of these chains as starting points for the harmonic identification and constrained stages of the search. The initial MHMC runs took 2–3 weeks each, the harmonic identification was then quick (a few hours), but the final constrained stage also took several weeks. We expect these run times to decrease as the algorithm is refined, but a final run time of ∼1 week is a reasonable guess. We should mention that the search of Challenge 1B.3.2 was the most advanced at the time of submission, and this was the only source for which we had begun the constrained jump phase of the search by that time.

A harmonic analysis indicated that we were in the vicinity of secondary maxima for sources 1B.3.1, 1B.3.3, 1B.3.4 and 1B.3.5. The harmonics of our best guess parameters for 1B.3.2 appeared to be approximately in the right place, although the total SNR of this point did not lie within the MLDC prior, indicating that the parameters were not completely correct. The submitted results are summarized in tables 2 and 3, along with the true values of the Challenge parameters.

As expected, our results for 1B.3.2 agreed quite well with the true parameters, but in the other cases we were quite far off. The recovered parameters correspond to secondaries of the likelihood in all cases, which share harmonics with the true signal. For all sources except 1B.3.1, we appeared to have locked onto the antipodal sky location, despite the inclusion of a sky position flip as one of the proposal distributions. We hope to avoid this in the future by implementing *delayed rejection* [18] which will be discussed in section 4.

4. Summary and future plans

Our final search algorithm for identifying isolated EMRIs in instrumental noise was as follows: (1) run a set of chains without annealing for a small number of chain steps; (2) identify

Table 3. Comparison of best recovered parameters at time of MLDC submission to true parameters—sky position and source orientation.

	Source	β	ϕ_S	θ_K	ϕ_K
1B.3.1	True	0.552 58	4.9104	1.7625	2.0472
	Recovered	0.645 58	4.9473	2.2005	1.7141
1B.3.2	True	0.359 70	4.6826	3.0372	4.0382
	Recovered	−0.874 34	1.1145	1.5836	6.0901
1B.3.3	True	0.981 66	0.709 67	1.5364	4.1487
	Recovered	−0.694 82	3.8708	2.2751	3.8650
1B.3.4	True	−0.980 20	0.978 73	1.7601	4.1164
	Recovered	0.813 38	3.2806	0.609 80	6.0855
1B.3.5	True	−1.1092	1.0876	0.840 39	5.7564
	Recovered	0.390 53	1.9690	1.9561	5.1293

harmonics of the true signal from the highest SNR points found in the preliminary chains; (3) run a partially constrained search, including temperature and time annealing, to refine the source parameters. The search runs were incomplete at the time of the MLDC deadline, but subsequent runs have moved closer to what we now know are the true parameters. The convergence rate is slow, however, so we are currently exploring several improvements to the algorithm, including

- *Fast waveform model.* By using the low-frequency approximation to the detector response, and interpolation of the barycentre waveform, it is possible to speed up the waveform and FIM evaluations which are the bottlenecks in the current code.
- *Semi-coherent analysis.* As considered for template-based searches in [1], by dividing the data stream into sections of a few months in length, and searching these separately in parallel, the search speed is increased. The main difficulty is forcing parameter consistency between the different segments.
- *Delayed rejection.* This is a technique whereby the chain is forced to accept a (large) jump in the parameter space, but the likelihood at the original point, $\vec{x}_i$, is recorded. The chain is then allowed to run (with small jumps) for a pre-specified number of steps, e.g., 100, to reach a final point, $\vec{x}_f$, before the Metropolis–Hastings ratio is evaluated using $\vec{x}_i$ and $\vec{x}_f$. If this step is rejected, the chain goes back to $\vec{x}_i$ and another jump is proposed. This technique can help the chain jump to widely separated secondaries in the parameter space, and it might also encourage the chain to switch to the correct antipodal sky position, as mentioned earlier.
- *Numerical F-statistic.* It is relatively straightforward to construct a proposal that changes the intrinsic parameters of the source in order to achieve a rotation of the signal harmonics, i.e., to keep the dominant harmonic in the same place, but change the indices (n, l, k) of that harmonic. However, it will usually be necessary to change the extrinsic parameters and phase angles in order to find a high SNR at the new point. This can be accomplished by proposing a jump in the intrinsic parameters, and then finding the best-fit extrinsic parameters at the new point, either by using a template grid or by using a mini ($\sim$100 iteration) Monte Carlo chain that explores only the extrinsic parameters. The likelihood maximized over extrinsic parameters can then be used to evaluate the Metropolis–Hastings ratio for the proposed jump in intrinsic parameters.

A short-coming of the algorithm described here will become apparent when the data stream becomes more complex and contains multiple sources, as in MLDC Round 3. When multiple sources are present, distinguishing between a weak harmonic that is a sideband of an identified bright source and one that is the dominant harmonic of a second, weaker signal, is difficult. The location in time and frequency, and the shape of the track, will be useful diagnostics, but some confusion will be inevitable. As an alternative to identifying harmonics to use as constraints, an understanding of how harmonic rotations/shifts can be achieved by parameter changes should allow the formulation of proposal distributions that will move the chain from one secondary to another (similar to the 'island hopping' used in supermassive black hole binary searches [15]). The aim of such a proposal would be to make the chain jump between points in parameter space that had harmonics in common with the true signal and hence the chain should find the true parameters more quickly. Such proposals will be explored when we apply these algorithms to the Round 3 data.

Acknowledgments

References

[1] Gair J R, Barack L, Creighton T, Cutler C, Larson S L, Phinney E S and Vallisneri M 2004 *Class. Quantum Grav.* **21** S1595
[2] Wen L and Gair J R 2005 *Class. Quantum Grav.* **22** S445
[3] Gair J R and Wen L 2005 *Class. Quantum Grav.* **22** S1359
[4] Wen L, Chen Y and Gair J R 2006 *AIP Conf. Proc.* **873** 595
[5] Gair J R and Jones G 2007 *Class. Quantum Grav.* **24** 1145
[6] Gair J R, Mandel I and Wen L 2007 *J. Phys.: Conf. Ser.* **122** 012037 (*Preprint* arXiv:0710.5250)
[7] Andrieu C and Doucet A 1999 *IEEE Trans. Signal Process.* **47** 2667
[8] Umstatter R, Christensen N, Hendry M, Meyer R, Simha V, Veitch J, Vigeland S and Woan G 2005 *Phys. Rev. D* **72** 022001
[9] Cornish N J and Crowder J 2005 *Phys. Rev. D* **72** 043005
[10] Crowder J and Cornish N J 2007 *Phys. Rev. D* **75** 043008
[11] Cornish N J and Porter E K 2006 *Class. Quantum Grav.* **23** S761
[12] Wickham E D L, Stroeer A and Vecchio A 2006 *Class. Quantum Grav.* **23** S819
[13] Cornish N J and Porter E K 2007 *Phys. Rev. D* **75** 021301
[14] Cornish N J and Porter E K 2007 *Class. Quantum Grav.* **24** S501
[15] Cornish N J and Porter E K 2007 *Class. Quantum Grav.* **24** 5729
[16] Stroeer A, Gair J R and Vecchio A 2006 *AIP Conf. Proc.* **873** 444
[17] Arnaud K A *et al* 2006 *AIP Conf. Proc.* **873** 619
[18] Cornish N J 2008 *Preprint* arXiv:0804.3323
[19] Babak S *et al* 2008 *Class. Quantum Grav.* **25** 114037 (*Preprint* arXiv:0711.2667)
[20] Babak S *et al* 2008 *Class. Quantum Grav.* **25** 184026 (*Preprint* arXiv:0806.2110)
[21] Barack L and Cutler C 2004 *Phys. Rev. D* **69** 082005