Detecting and Characterizing EMRIs in the second round of Mock LISA Data Challenges

Neil J. Cornish
Department of Physics, Montana State University, Bozeman, MT 59717

This brief note describes a search for EMRI signals in the MLD C Challenge 1.3.1 and 1.3.2 data sets using a Metropolis-Hastings Monte Carlo (MHMC) search algorithm. The algorithm was tested on the 1.3.1 training data, and it was able to detect the signal and recover the source parameters to a level consistent with the Markov Chain Monte Carlo (MCMC) derived parameter errors. The tests also indicated that the EMRI posterior distribution function has a very complex topography, with vast numbers of secondary maxima both near and far from the injected parameter values. Very good fits to the data could be achieved with parameters very different from those used to inject the signal. Once our search algorithm locked onto one of these secondary maxima there was little chance that it would later transition to the primary maximum. The MHMC algorithm used to analyze the challenge data sets is far from optimal, and it is hoped that later versions will be more successful.

I. HUNTING EMRIS

The Montana Gravitational Wave Astronomy Group has developed a variety of Metropolis-Hastings Monte Carlo (MHMC) search algorithms for use in LISA data analysis. The MHMC approach is closely related to the Markov Chain Monte Carlo (MCMC) method for mapping posterior distribution functions (PDF). The main difference is that the MHMC search chains do not have to be reversible, which opens up the possibility of using techniques such as simulated annealing, and transition rules that employ the past history of the chain. The MHMC approach has been successfully applied to searches for simulated signals from galactic binaries [1] and massive black hole binaries [2].

The MHMC algorithm used in the EMRI searches shares many features in common with the MBHB and galactic binary (GB) algorithms. The algorithm employs a mixture of transition proposals: large and small jumps in one or all of the parameters based on uniform draws from the parameter priors; and normal jumps along local eigen-directions of the Fisher Information Matrix. The current version of the algorithm does not employ any of the specially tailored “island hopping” jumps that promote movement between widely separated maxima of the PDF. It is likely that the inclusion of island hopping transitions will significantly improve the performance of the search (if our experience with MBHBs and galactic binaries is anything to go on). In contrast to the GB and MBHB search algorithms, the EMRI search only updated the Fisher Information Matrix (FIM) once every 1000 steps. This was done as the computation of the FIM requires 24 calls to the waveform generation routine, and these calls are better spent in advancing the search chain. The frequency of the FIM updates was taken as a trade off between the cost of generating the FIM, and the improvement in the chain transition rate that occurs when the FIM is updated. As with all of our MHMC algorithms, simulated annealing was used to ensure complete coverage of the parameter search range during the initial phase of the search.

The main new ingredient in the EMRI search was a hierarchical procedure that started out using small segments of the full data stream and gradually built up to using the full signal. In that respect the search shares some features in common with the method proposed in Ref. [3]. The first stage of the search divided the two-year data stream into 32 segments (about 23 days each). The best fit(s) from the first stage were then used to start the search at the next level (16 segments) and so on. At each transition simulated annealing was used to promote movement of the chains. The segmented approach has many advantages: first, it takes a lot less time to generate the waveforms for the early phases of the search; and second, it is possible to get a good match to the signal with a wider range of parameters. For example, with a 23 day data segment the sky location is not well determined, so it is possible to match the signal with the wrong ecliptic latitude and longitude. On the other hand, the best fit parameters on a short data segment can be very far from the correct values, and this can trap the search in the wrong region of parameter space. In tests on the 1.3.1 training data it was often found that simulated annealing was unable to free the chain from these secondary maxima. Various remedies were tried, with the most successful being to use two data segments simultaneously, one from early in the data stream and one from later in the data stream. The orbital evolution was carried through coherently between the segments, so the signal was then pinned down at two places. This gave the search vital information about overall frequency evolution which is lacking when using a single short data segment. Parameters such as the mass of the central black hole were much better determined by this new approach. The two-segment approach strikes a balance between making the signals easier to detect by “widening the target zone” and returning
parameters that are far from the central maximum. The MHMC approach is ideally suited to a hierarchical search as it automatically ignores parameters that are poorly determined (jumps in these parameters directions are always accepted), and even a marginal signal detection is enough to move the search into the right general area of parameter space, which then enables the search to quickly latch on to the signal as more data is incorporated.

Even with a hierarchical approach the computational cost of an EMRI search was still out of reach for someone without access to a supercomputer. For example, the existing codes for generating the Barack-Cutler EMRI waveforms [4] took several minutes to run each set of parameters. The first step was to write a new code that employs a fast-slow decomposition of the waveform generation. This new code can generate a two-year waveforms in about a second, and tens of segment waveforms per second. The other cost saver was to use spectrograms to pre-select the best waveform segments to use in the various stages of the hierarchy (these bright segments also narrow down the sky location but this information was not used in the search). The spectrograms were also used to narrow the prior range on the initial orbital frequency, \( \nu_0 \). The various lines in the spectrogram correspond to the harmonics \( n\nu + 2f_\gamma + mf_\alpha \), where \( f_\gamma = \dot{\gamma}/(2\pi) \) and \( f_\alpha = \dot{\alpha}/(2\pi) \). Since \( f_\alpha < 0.1\nu \) the tracks come in bunches corresponding to a particular \( n \). Differentiating the frequencies of tracks in each bunch gives \( \nu \), though there is some ambiguity in pairing up tracks between bunches (they may be correspond to different \( m \)'s). Moreover, \( f_\gamma = f_\gamma + kAf_\alpha \), where \( k \) is any integer and \( A \) is a constant of order unity. This means that the sidebands are not always integer multiples of \( f_\alpha \) apart. Using the tracks it is possible to bound \( \nu_0 \) to a range of \( \pm 1e-6 \) Hz, with a possible ambiguity of \( \pm 1e-5 \) Hz due to pairing up the wrong sidebands. This can be compared to the prior range on \( \nu_0 \) set by the priors on the masses and eccentricity at plunge, which restricts \( \nu_0 \) to be in the range \( 1.7e-4 \) \( \rightarrow 2e-4 \) Hz for Challenge 1.3.1. In other words, the allowed range for \( \nu_0 \) is already just \( 3e-5 \) Hz, so information from the tracks is not much help unless you gamble on having paired up the right sidebands (which we did). It should be emphasized that the tracks were just used to reduce the computational burden by a factor or 10 or so, and with more computer resources the spectrograms would not be needed.

The results reported for 1.3.1 should give a good match to the injected signal. It is still quite likely that the parameters values correspond to one of the many secondary maxima. With more runs (or an improved algorithm) it would have been possible to find the central maximum, just as was done with the training data. The chains were still being run in a “finisher” mode when the submission deadline came around, and the convergence was not complete. Full MCMC posterior distributions will be available in the days following the submission deadline. The results reported for 1.3.2 will give a so-so match to the injected signal. The initial search had to be cut short and several stages in the hierarchical refinement had to be skipped in an effort to meet the deadline. With only a laptop and a workstation to run on, computer resources were the major limitation. If a supercomputer had been used the current code would have been able to find all the EMRIs in 1.3.x and 2.2. The problem of secondary maxima can be overcome by a shotgun approach of running multiple chains and selecting the best runs for further refinement. A genetic algorithm will be used to carry out this selection procedure for runs on Challenge 1b.

Examples of the MCMC derived parameter posteriors for the training data sets are shown on the following pages.

ACKNOWLEDGMENTS

This work was supported by NASA Grant NNG05GI69G.

REFERENCES

FIG. 1: Marginalized PDFs for Training data set 1.3.1 are shown in red. The blue lines are the FIM estimates.
FIG. 2: Marginalized PDFs for Training data set 1.3.2 are shown in red. The blue lines are the FIM estimates.
FIG. 3: Marginalized PDFs for Training data set 1.3.3 are shown in red. The blue lines are the FIM estimates. Note the pronounced bimodality in the spin orientation parameters $\theta_K$, $\phi_K$, and the spin-orbit orientation parameters $\gamma$, $\alpha$. 
FIG. 4: Marginalized PDFs for training data set 1.3.4 are shown in red. The blue lines are the FIM estimates. In this example the FIM estimates for the intrinsic parameters are much better. In contrast to the other examples, the FIM estimate for the error in the mass of small BH appears to be too large.
FIG. 5: Marginalized PDFs for training data set 1.3.5 are shown in red. The blue lines are the FIM estimates. In this example the FIM estimates are all excellent.