Introduction

Extreme mass ratio inspirals (EMRIs) are one of the most exciting sources of gravitational waves (GWs) that laser interferometer space antenna (LISA) is expected to observe. These inspirals are formed when a stellar mass compact object (CO) with mass \( 1 - 10M_\odot \) is captured in a strong orbit of a spinning super massive black hole (SMBH) with mass \( \sim 10^5 - 10^6M_\odot \) and subsequently, under the influence of gravitational radiation, inspirals gradually into SMBH through the emission of GWs. The detection and characterization of signals from such sources will help to understand the structure and formation of SMBHs and the characteristics of the space-time around them, such as lensing-thirring or frame-dragging effects. Bayesian approach equipped with sophisticated Markov chain Monte Carlo (MCMC) algorithms has been used for GWs signal processing with great successes.

The Waveform Model and its Features

The EMRI signals given in MLDC releases are based on the so-called “analytic kludge” waveform (AKW) model \([1]\). Although these waveforms are not very faithful approximation, they still capture the main features of realistic EMRIs and are easy to generate. EMRI signals are very weak, have a very complicated spectrum, and are computationally expensive. Another version of AKW is the truncated AKW (TAKW) model given in \([2]\). TAKWs are \( \sim 3 \) times faster than the full AKWs and have a very good agreement with them, and therefore we generally use this model in our search codes. From a statistical point of view, the detection and parameter estimation for EMRI signals is a challenging problem. The parameter space is huge (14 parameters) and the likelihood surface is overwhelmed by local maxima. In practice, the ordinary MCMC algorithms fail to explore such multi-model surfaces efficiently and accurately.

LISA Response

We use LISA simulator \([3]\) data in which the LISA response to GWs is given in the form of three time delay interferometry (TDI) observables namely \( X, Y, Z \). These initial TDI observables are highly correlated. To orthogonalize them another set of observables namely \( A, E, T \) is created for actual data analysis \([4]\). In the past, in our MCMC codes, to compute the LISA response for our model polarization signals \( (h_+, h_\times) \) we used the actual LISA simulator but it was too slow. Now, we use low frequency approximation to approximate LISA response \([1, 5]\). Low frequency approximation is very effective and is considerably faster than LISA simulator (\( \sim 13 \) times in our case) and is also much more flexible than LISA simulator in the sense of parameter priors specifications, i.e. we can now use broader priors for signal parameters with faster MCMC follow up.

The Noise Spectrum

The LISA noise mainly has three components: instrument noise, confusion noise from galactic binaries and confusion noise from extra-galactic binaries. These confusion noises are expected to be dominating the LISA data. There have been proposals to identify and subtract such confusion sources \([6]\) but in general such a strategy may not be a feasible option because firstly it is time consuming as there will be numerous such sources; secondly subtraction of a poorly matched signal may destroy the data which may change the characteristics of other potential signals present in the same data. In reality, the noise properties will be unknown. Bayesian approach can be employed to model the unknown noise in a very logical and simple way. Instead of modeling the confusion sources individually, we can treat the whole melange as noise. In Fourier paradigm each ordinate of the power spectral density has a \( \chi^2 \)-distribution with 2 degrees of freedom. Using the scaled \( \ln \chi^2 \) distribution as prior for spectral ordinates, a posterior, which is also an \( \ln \chi^2 \) distribution can be established. This posterior distribution can be updated using the current residuals based on the updated signal parameters, through a simple Gibbs step without any additional computing costs \([7]\).

The Whittle Likelihood Function

The likelihood function which incorporates the noise spectrum as an unknown quantity is the Whittle’s likelihood function \([7, 8]\) defined as

\[
\mathcal{L}(\theta) = K \times \exp \left[ -\sum_{j=1}^{n} \log(S_j(f_j; \theta)) + \frac{\left| \tilde{z}(f_j) - \tilde{z}(f_j; \theta) \right|^2}{S_j(f_j; \theta)} \right]
\]

where \( \theta \) are signal parameters, \( f_j \) the Fourier frequencies and \( \tilde{z}(f_j; \theta) \) the examined frequency range, \( \tilde{z}(f_j) \) and \( S_j(f_j; \theta) \) the Fourier transformed observables and model respectively, \( S_j(f_j; \theta) \) the unknown one sided power spectral density depending on \( \theta \) and \( K \) the normalizing constant.

Parallel Tempering MCMC

Parallel tempering MCMC (PTMCMC) is a powerful twist of traditional Metropolis-(Hastings) algorithm which is very effective in improving the mixing and particularly in escaping the local maxima. The algorithm works by running multiple MCMC chains in parallel. Each chain simulates a separate target density characterized by a different temperature and occasionally attempts swap of its current state with other chains. In principle a high temperature chain sees the density of the target distribution as more flattened relative to a low temperature chain, which means that high temperature chain can move more freely across the valleys of low probability regions in between of local modes. Swaps of states between chains enable the chains stuck at local maxima to get released and hence move more efficiently towards the global maximum.

Applications and Results

We have applied our algorithm to several MLDC datasets containing single and multiple EMRI sources and have obtained very good results. Some preliminary results obtained by using a long week long data segments from MLDC Round 4 training and blind datasets are given in the following figures.

Conclusion

Our preliminary results show that Bayesian approach with PTMCMC algorithm has a great promise for conducting efficient EMRI searches and estimation of their parameters. Simultaneous estimation of signal parameters as well as the noise spectrum will be inevitable in realistic situations and this can be effectively done in Bayesian framework. We are conducting more runs on MLDC Round 4 datasets and we hope to deliver much improved results in the near future. We hope that our algorithm will be very useful in realistic situations.

References

\[3\] N J Cornish et al. www.physica.montana.edu/LISA.