

Solution to the galactic foreground problem for LISAJeff Crowder^{1,2} and Neil J. Cornish¹¹*Department of Physics, Montana State University, Bozeman, Montana 59717, USA*²*Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109, USA*

(Received 27 November 2006; published 23 February 2007)

Low frequency gravitational wave detectors, such as the Laser Interferometer Space Antenna (LISA), will have to contend with large foregrounds produced by millions of compact galactic binaries in our galaxy. While these galactic signals are interesting in their own right, the unresolved component can obscure other sources. The science yield for the LISA mission can be improved if the brighter and more isolated foreground sources can be identified and regressed from the data. Since the signals overlap with one another, we are faced with a “cocktail party” problem of picking out individual conversations in a crowded room. Here we present and implement an end-to-end solution to the galactic foreground problem that is able to resolve tens of thousands of sources from across the LISA band. Our algorithm employs a variant of the Markov chain Monte Carlo (MCMC) method, which we call the blocked annealed Metropolis-Hastings (BAM) algorithm. Following a description of the algorithm and its implementation, we give several examples ranging from searches for a single source to searches for hundreds of overlapping sources. Our examples include data sets from the first round of mock LISA data challenges.

DOI: [10.1103/PhysRevD.75.043008](https://doi.org/10.1103/PhysRevD.75.043008)

PACS numbers: 95.55.Ym, 04.80.Nn, 95.85.Sz

I. INTRODUCTION

Galactic compact binary systems are expected to be the major source of gravitational waves detected by the LISA observatory [1]. Tens of millions of such binaries will be emitting gravitational waves in the LISA band [2–6]. Though most will have a signal-to-noise ratio (SNR) too low to be detectable, tens of thousands are expected to be resolvable if optimal signal analysis algorithms are available [7]. The signals from unresolved binary systems will constitute a gravitational wave confusion noise. Identification of the brighter galactic binaries will be an important aid to resolving other sources for LISA, such as supermassive black hole binaries (SMBHBs) [8–12] and extreme mass ratio inspirals of compact objects into supermassive black holes (EMRIs) [13–16]. The SMBHB, EMRI, and compact binary signals have small, but non-vanishing overlap with one another, so we will ultimately want a simultaneous fit to all sources and source types.

Various techniques have been proposed to extract the parameters of sources from the LISA data stream. The methods include Markov chain Monte Carlo methods [17–21], genetic algorithms [22], iterative methods [23,24], grid-based template searching [25], tomographic reconstruction [26], and time-frequency methods [27]. For most of these methods (tomography and time-frequency analysis being the exceptions), optimal filtering is accomplished through the construction of templates describing the signals from all sources in the data stream. For LISA, the vast number of sources involved makes a direct approach (such as a grid-based template bank) computationally impractical. It is for this reason ergodic methods such as Markov chain Monte Carlo (MCMC) and genetic algorithms have been applied to the LISA data analysis problem.

In this work we develop an extension of the MCMC method [28–30] that is able to search the entire LISA band and simultaneously solve for thousands of galactic binaries. Our approach is based on the observation that, while some signals can have significant overlap, signals that are well separated in frequency have little or no overlap [31]. We exploit this quasilocality by breaking the search up into subregions in frequency, taking care with edge effects. In a departure from our previous approach [17], we do not try and update all the source parameters simultaneously at each iteration. This greatly reduces the computational cost, while still providing a global solution as the full multitemplate is used to evaluate the likelihood. The multitemplate is the superposition of all the individual 7 parameter search templates. We refer to the individual components of the multitemplate as “searchers” as the evolution of the parameter chains is reminiscent of bloodhounds latching onto a scent trail. The parameter updates are done in small blocks of highly correlated search templates, and the solution is updated using Metropolis-Hastings sampling. Simulated annealing is employed during the search phase to improve the mixing of the chains. We demonstrate this “blocked annealed Metropolis-Hastings” (BAM) algorithm on simulated LISA data that contains the signals from monochromatic white dwarf binary systems (WD-WD) immersed in Gaussian instrumental noise.

The paper is organized as follows: In Sec. II we review the LISA galactic foreground problem, and give a brief introduction to the MCMC algorithm and the Metropolis-Hastings sampling kernel. Additionally, we give a description of the BAM algorithm and demonstrate how its computational cost scales with the number of sources being searched for. Section III discusses various aspects of using the BAM algorithm, such as hierarchical searching, opti-

mal choices for search ranges, stopping criteria, and rates of occurrence for false positives and false negatives. Example searches are performed in Sec. IV, ranging from individual sources to hundreds of sources in a restricted range of the LISA band. Concluding remarks are made in Sec. V.

II. ALGORITHM OVERVIEW

In this section we describe the BAM algorithm and the issues and difficulties that we worked through in the development process. We start with a brief review of the basic MCMC approach and Metropolis-Hastings sampling. This is followed by a review of how the extrinsic source parameters are analytically solved for using the F -statistic. We then describe how the speed of the BAM algorithm scales with the number of sources that are being searched for.

A. The galactic foreground problem

It is anticipated that tens of millions of compact binary systems will emit radiation at frequencies that lie within the LISA band. The majority of these systems will be detached double white dwarf binaries, or interacting double white dwarf binaries of the AMCVn class. A much smaller fraction will involve naked helium stars, neutron stars, or stellar mass black holes. The number of compact binaries is predicted to increase sharply with decreasing frequency, and it has been estimated that the average number of systems per $1/\text{year} \approx 31$ nHz frequency bin exceeds unity for frequencies below ~ 3 mHz [4–6]. This inevitably leads to a confusion limited residual at low frequencies, and is the reason why most of the resolvable systems have frequencies in the 2 to 5 mHz range [7].

It has been estimated that it will be possible to resolve ~ 15 000 individual compact binary systems from a 1-year

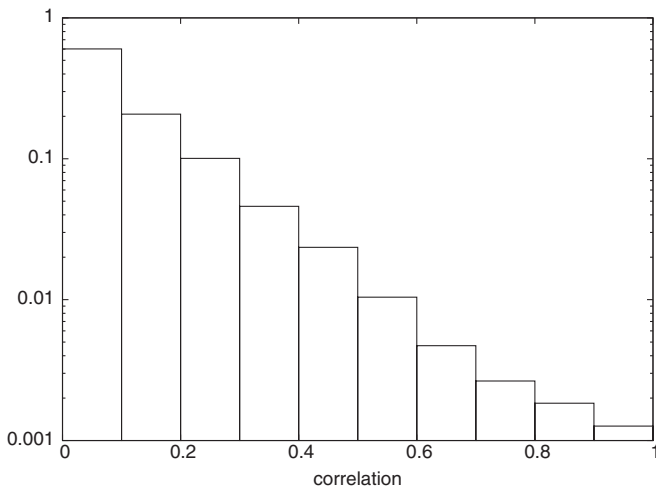


FIG. 1. Histogram of the largest correlations between bright galactic binaries (SNR > 10).

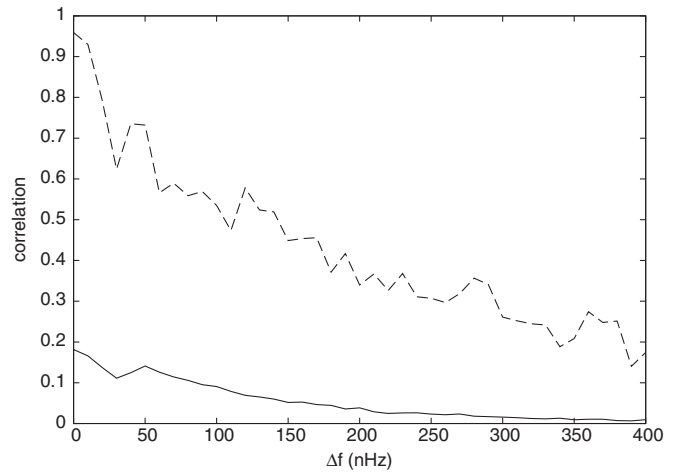


FIG. 2. The maximum (dashed line) and average (solid line) correlation between bright galactic sources as a function of the difference in their barycenter frequencies (binned in units of 10 nHz).

LISA data stream [7]. Since the signals are continuous, and described by $7 \rightarrow 9$ parameters, it might appear that we are faced with the problem of having to simultaneously solve for a model with $\sim 10^5$ components. However, the bandwidth of each individual signal is limited ($\sim 0.1 \rightarrow 10$ μHz), so most of the signals have very small overlap. Taking the ~ 17 000 binary systems of the Nelemans-Yungelson-Zwart [6] galaxy model that were identified as having signal-to-noise ratios > 10 by Timpano *et al.* [7] and computing the maximum correlation between them yields the correlation histogram shown in Fig. 1. We find that 68% of the bright sources suffer correlations no larger than 0.2 with other bright sources. Here the correlation is defined as the absolute value of the overlap, where the overlap between signals a and b is defined as

$$C_{ab} = \frac{|a|b|}{(|a|)^{1/2}(|b|)^{1/2}}, \quad (1)$$

where $|a|b|$ denotes the noise-weighted inner product

$$(a|b) = 2 \int_0^\infty \frac{\tilde{a}^*(f)\tilde{b}(f) + \tilde{a}(f)\tilde{b}^*(f)}{S_n(f)} df, \quad (2)$$

and $S_n(f)$ is the one-sided noise spectral density. Figure 2 shows how the correlation between bright sources decays as they get further apart in frequency, and underlines the fact that the problem of finding a global fit to the galactic foreground is an essentially local problem in frequency space.

B. The Markov chain Monte Carlo algorithm and Metropolis-Hastings sampling

The MCMC algorithm is becoming a familiar tool in gravitational wave data analysis. Initially introduced to the field by Christensen and Meyer [32], its application to

ground-based interferometers has been explored in the context of parameter extraction of coalescing binaries [33] and spinning neutron stars [34]. With its ability to explore large parameter spaces while simultaneously performing model selection and noise estimation, the MCMC method is ideally suited to the LISA data analysis problem. The reverse-jump MCMC algorithm has been applied to the LISA-like problem of identifying a large, yet unknown number of sinusoids in simulated Gaussian noise [35,36]. It was shown that the method could correctly identify the number of resolvable signals present in the data and recover the signal parameters and an estimate of the noise level. The MCMC approach was first applied to simulated LISA data in the context of galactic binaries [17], and it has since been applied to SMBHBs [18–20].

In using an MCMC approach, one wants to generate a sample set $\{\vec{x}\}$ that corresponds to draws made from the posterior distribution of the system $p(\vec{x}|s)$. The procedure for generating such a set is surprisingly simple.

We begin at a point in the parameter space of the binary system(s), \vec{x} (which may or may not be chosen at random), and propose a jump to a new position, \vec{y} , based on some proposal distribution, $q(\cdot|\vec{x})$. The Hastings ratio is calculated using

$$H = \frac{p(\vec{y})p(s|\vec{y})q(\vec{x}|\vec{y})}{p(\vec{x})p(s|\vec{x})q(\vec{y}|\vec{x})}, \quad (3)$$

where $p(\vec{x})$ is the prior of the parameters at \vec{x} , $q(\vec{x}|\vec{y})$ is the value of the proposal distribution for a jump from \vec{x} to \vec{y} , and $p(s|\vec{x})$ is the likelihood at \vec{x} . The likelihood function, if the noise is a normal process with zero mean, is given by [37]

$$p(s|\vec{x}) = C \exp[-\frac{1}{2}((s - h(\vec{x}))|(s - h(\vec{x})))], \quad (4)$$

where the normalization constant C is independent of the signal, s . The jump will be accepted with probability $\alpha = \min(1, H)$. If the jump is rejected (Metropolis rejection [28]) the chain remains at its current state, \vec{x} . Repeated jumps will produce a Markov chain whose stationary distribution is equal to the posterior distribution in question, $p(\vec{x}|s)$. Andrieu *et al.* [35,38] provide a more general and thorough review of MCMC methods.

The convergence to the correct posterior will occur for any (nontrivial) proposal distribution [39]. However, the closer the proposal distribution is to the posterior distribution, the quicker the chain will converge. Since we do not know the form of the posterior in advance of running the chain, we instead opt for flexibility in our choice of proposal distribution. We accomplish this by using a mixture of proposal distributions, including occasional “bold” proposals that attempt large changes in the parameter values along with many “timid” proposals that attempt small changes in the parameter values of the chain (for a detailed description of some of these proposals see [17]). Also added to our list of proposals are a few “tailored”

proposals. These are proposed jumps based on our knowledge of the symmetries and degeneracies of the likelihood surface. For example, for LISA there is a secondary maxima in the likelihood surface of a binary system that corresponds to a reflection about the ecliptic equator combined with a shift of π in the ecliptic longitude. Thus, we include a proposal that attempts such a reflection.

Another tailored proposal that was highly effective uses the fact that local maxima occur in the likelihood surface at multiples of the modulation frequency ($f_m = 1/\text{year}$). The presence of these maxima are easily understood. The detector response imparts sidebands on the monochromatic barycenter signal to produce a comb whose teeth are spaced by the modulation frequency. A template with a Barycenter frequency offset from the signal by a multiple of f_m will produce a similar comb whose teeth align with those of the signal. The fit can be improved by adjusting the other template parameters to better match the shape of the source comb. Figure 3 shows the power spectrum of a source and a template offset by $1f_m$ in frequency, while in Fig. 4 the template has had its other parameters altered to better fit the source comb. This improves the log likelihood from 16.7% to 66.4% of the true parameter value. It proved to be difficult to find an analytic description of how the other parameters should be modified to improve the fit, so we went with a simple proposal that shifted the frequency of the source by $1f_m + \epsilon$, and shifted the other parameters using a uniform draw in a small range around the current binary parameters.

Since the search used an F -statistic to extremize over the amplitude, inclination angle, polarization angle, and initial phase, we only needed to assign priors to the frequency and sky location. For the frequency we used a uniform prior across the search range, and for the sky locations we used a distance weighted galactic distribution.

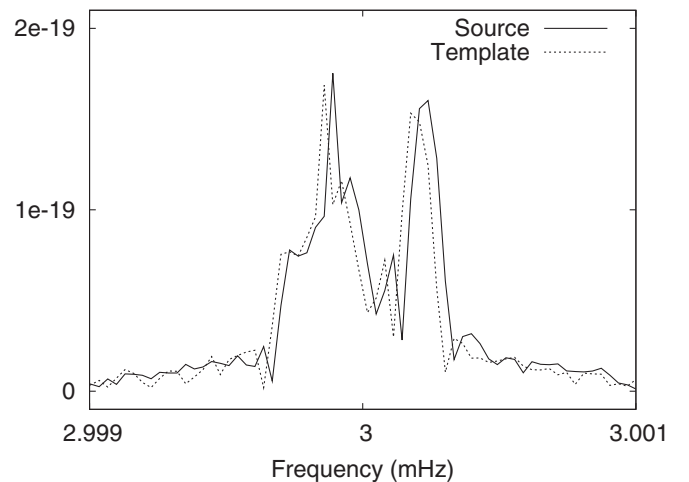


FIG. 3. Spectrum of a source and template with the same parameters save their frequencies, which differ by one modulation frequency. The source has $f = 3.0$ mHz, $A = 1.4 \times 10^{-22}$, $\theta = 1.0$, $\phi = 3.14159$, $\psi = 0.5$, $\iota = 0.785398$, and $\varphi_o = 0.0$.

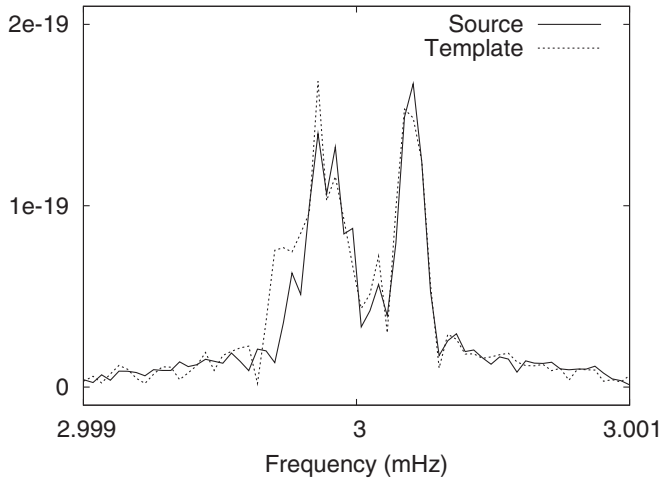


FIG. 4. Spectrum of a source and a template with frequencies that differ by one modulation frequency. The parameters of the template have been adjusted to maximize the overlap. The source has the same parameters as in Fig. 1, while the template has $A = 1.47254 \times 10^{-22}$, $f = 3.000031688$ mHz, $\theta = 0.787087$, $\phi = 3.18781$, $\psi = 0.51218$, $\iota = 0.89419$, and $\varphi_o = 0.179034$.

While a general MCMC algorithm functions as both a means to search the likelihood surface and sample from the posterior distribution, the Markovian nature is only required for the sampling portion of the process. For the purpose of searching to determine the neighborhood of the true parameters, which is the focus of this work, we allow ourselves the extra freedom provided by non-Markovian Metropolis-Hastings sampling and simulated annealing (described below). Once the search phase is complete the non-Markovian moves are suspended, and the algorithm performs a standard MCMC exploration of the posterior distribution. One example of a non-Markovian proposal involves how we implement the previously described $1f_m$ jump. Early in the search phase, if one of these proposals was accepted, a second (identical) jump was attempted. The reason for this is that just as a local maximum occurs $\sim 1f_m$ away from the global maximum, another (smaller) local maximum occurs $\sim 2f_m$ away. In fact, a chain of maxima occur spaced about one modulation frequency apart from each other in likelihood space (see Fig. 5). This non-Markovian move allowed searchers to “island-hop” around the likelihood surface if it found itself on one of the maxima in the island chain.

To encourage the chains to explore the full parameter space, and to discourage the chains from getting stuck on local maxima, we employ simulated annealing during the search phase. This is done by multiplying the noise-weighted inner product by an “inverse temperature” β , and applying the cooling schedule

$$\beta = \begin{cases} \beta_0 (\frac{1}{\beta_0})^{i/N_c} & 0 \leq i \leq N_c \\ 1 & i > N_c. \end{cases} \quad (5)$$

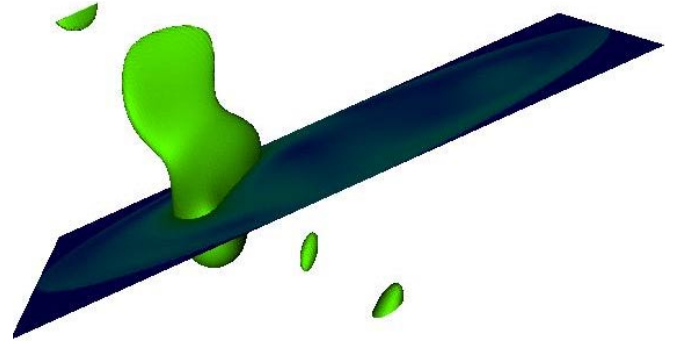


FIG. 5 (color online). Contours of constant likelihood for a single, monochromatic WD-WD binary system. Sky position directions lie in the plane shown, while frequency is directed orthogonal to the plane. The visible maxima are separated by approximately one modulation frequency.

Here β_0 is the initial heating factor and N_c is the length of the annealing phase. The choice of β_0 depends on the SNR of the sources. When very bright sources are present we set $\beta_0 \sim 10^{-2}$, while smaller values of $\beta_0 \sim 10^{-1}$ work better once the brightest sources have been removed. Our criteria for choosing β_0 was that the chains should explore the full parameter range during the first 1000 or so iterations. If full range movement was not seen we dialed up the heat. The choice of N_c depends on β_0 , and can be set by demanding that it takes, say, 500 steps for β to decrease by a factor of 2.

C. F -statistic

The F -statistic [40] uses multiple linear filters to obtain the extremum for the likelihood over the extrinsic parameters of the signal. Using the F -statistic, one can search the intrinsic parameters, and recover the extrinsic parameters as a last step in the process. For a monochromatic binary detected by LISA, this reduces the search space to three parameters: frequency and sky location (θ , ϕ). Note: the tumbling orbit of LISA induces modulations of the frequency into the signal, thus creating an interdependency in the signal between frequency and sky location that prevents θ and ϕ from being treated as purely extrinsic quantities.

At low frequencies, the LISA response to a gravitational wave can be written

$$h(t) = h_+(t)F^+(t) + h_\times(t)F^\times(t), \quad (6)$$

where $h_+(t)$ and $h_\times(t)$ are the two polarizations of the incident gravitational wave [41,42], which for a monochromatic binary, to leading post-Newtonian order, are given by

$$\begin{aligned} h_+(t) &= A(1 + \cos^2 \iota) \cos(\Phi(t) + \varphi_0) \\ h_\times(t) &= -2A \cos \iota \sin(\Phi(t) + \varphi_0). \end{aligned} \quad (7)$$

$F^+(t)$ and $F^\times(t)$ are the beam pattern factors:

$$\begin{aligned} F^+(t) &= \frac{1}{2}(\cos 2\psi D^+(t) - \sin 2\psi D^\times(t)) \\ F^\times(t) &= \frac{1}{2}(\sin 2\psi D^+(t) + \cos 2\psi D^\times(t)), \end{aligned} \quad (8)$$

where $D^+(t)$ and $D^\times(t)$ are the detector pattern functions, whose mathematical form can be found in Eqs. (36) and (37) of Ref. [43].

In the gravitational wave phase

$$\Phi(t; f, \theta, \phi) = 2\pi f t + 2\pi f \text{AU} \sin\theta \cos(2\pi f_m t - \phi), \quad (9)$$

one can see the coupling of the sky location and the frequency through the term that depends on the radius of LISA's orbit, 1 AU, and its orbital modulation frequency, $f_m = 1/\text{year}$. For the low frequency galactic sources we are considering, the gravitational wave amplitude, A , is effectively constant. Thus (6) can be rearranged as

$$h(t) = \sum_{i=1}^4 a_i(A, \psi, \iota, \varphi_0) A^i(t; f, \theta, \phi), \quad (10)$$

where the time-independent amplitudes a_i are given by

$$\begin{aligned} a_1 &= \frac{A}{2}((1 + \cos^2 \iota) \cos \varphi_0 \cos 2\psi - 2 \cos \iota \sin \varphi_0 \sin 2\psi), \\ a_2 &= -\frac{A}{2}(2 \cos \iota \sin \varphi_0 \cos 2\psi + (1 + \cos^2 \iota) \cos \varphi_0 \sin 2\psi), \\ a_3 &= -\frac{A}{2}(2 \cos \iota \cos \varphi_0 \sin 2\psi + (1 + \cos^2 \iota) \sin \varphi_0 \cos 2\psi), \\ a_4 &= \frac{A}{2}((1 + \cos^2 \iota) \sin \varphi_0 \sin 2\psi - 2 \cos \iota \cos \varphi_0 \cos 2\psi), \end{aligned} \quad (11)$$

and the time-dependent functions $A^i(t)$ are given by

$$\begin{aligned} A^1(t) &= D^+(t; \theta, \phi) \cos \Phi(t; f, \theta, \phi) \\ A^2(t) &= D^\times(t; \theta, \phi) \cos \Phi(t; f, \theta, \phi) \\ A^3(t) &= D^+(t; \theta, \phi) \sin \Phi(t; f, \theta, \phi) \\ A^4(t) &= D^\times(t; \theta, \phi) \sin \Phi(t; f, \theta, \phi). \end{aligned} \quad (12)$$

Writing LISA's signal as a superposition of gravitational waves and noise,

$$s_\alpha(t) = h_\alpha(t, \vec{\lambda}) + n_\alpha(t) = \sum_{i=1}^N h_\alpha^i(t, \vec{\lambda}_i) + n_\alpha(t), \quad (13)$$

and defining the four constants $N^i = (s|A^i)$ and the M -matrix, $M^{ij} = (A^i|A^j)$, one can express (10) as a matrix equation:

$$M_{ij} a^j = N_i. \quad (14)$$

Therefore, with a signal from LISA and the three intrinsic parameter values, one can solve (by iteration or inversion) for the amplitudes, and thus the extrinsic parameters. The values of the intrinsic parameters are found by use of iterative Metropolis-Hastings sampling. The extrinsic pa-

rameter values are given by

$$\begin{aligned} A &= \frac{A_+ + \sqrt{A_+^2 - A_\times^2}}{2} \\ \psi &= \frac{1}{2} \arctan\left(\frac{A_+ a_4 - A_\times a_1}{-(A_\times a_2 + A_+ a_3)}\right) \\ \iota &= \arccos\left(\frac{-A_\times}{A_+ + \sqrt{A_+^2 - A_\times^2}}\right) \\ \varphi_0 &= \arctan\left(\frac{c(A_+ a_4 - A_\times a_1)}{-c(A_\times a_3 + A_+ a_2)}\right), \end{aligned} \quad (15)$$

where

$$\begin{aligned} A_+ &= \sqrt{(a_1 + a_4)^2 + (a_2 - a_3)^2} \\ &\quad + \sqrt{(a_1 - a_4)^2 + (a_2 + a_3)^2} \\ A_\times &= \sqrt{(a_1 + a_4)^2 + (a_2 - a_3)^2} \\ &\quad - \sqrt{(a_1 - a_4)^2 + (a_2 + a_3)^2} \\ c &= \text{sign}(\sin(2\psi)). \end{aligned} \quad (16)$$

This description of the F -statistic automatically incorporates the two independent LISA channels through the use of the dual-channel noise-weighted inner product.

We generalize the F -statistic to handle N overlapping sources by writing $i = 4K + l$, where K labels the source and $l = 1 \rightarrow 4$ labels the four filters for each source. The F -statistic has the same form, but with $4N$ linear filters N^i , and M^{ij} is a $4N \times 4N$ dimensional matrix. For slowly evolving galactic binaries, which dominate the confusion problem, the limited bandwidth of each individual signal means that the M^{ij} is band diagonal, which lessens the difficulty in solving (14) for the large numbers of sources that are expected to be detected.

D. The blocked annealed Metropolis-Hastings algorithm

Metropolis [28] noted that a multiparameter model can be explored by updating just one parameter at each iteration of the chain. A popular example of this approach is the Gibbs sampler [44], where each parameter is updated sequentially using a proposal distribution that is the conditional of the posterior evaluated at the current values of the other parameters. One drawback of single parameter updates is that correlations between parameters can hinder the mixing of the chain. The Blocked-Gibbs algorithm [45] addresses this problem by simultaneously updating blocks of highly correlated parameters. For our search algorithm, we adopt a similar strategy, and perform Metropolis-Hastings updates to blocks of searchers that are close in frequency.

The blocks in the BAM algorithm are small subunits of the frequency range being searched. As can be seen in

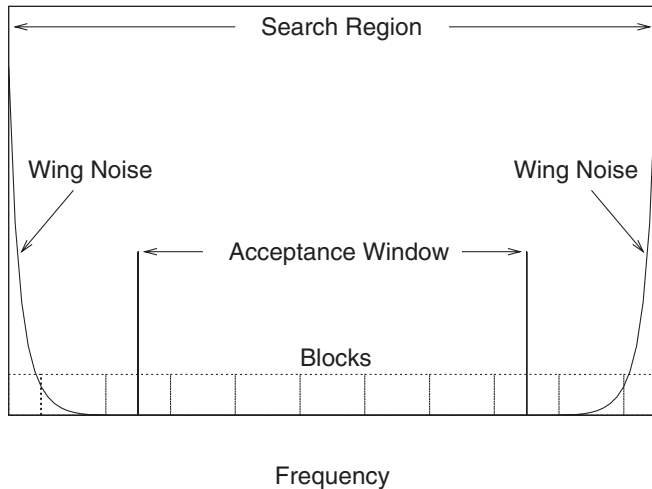


FIG. 6. A schematic representation of a search region in frequency space, showing the block structure of the BAM algorithm.

Fig. 6, which shows a schematic representation of a search region in our BAM algorithm, the search region is broken up into equal sized blocks. The algorithm steps through these blocks sequentially, updating all sources within a given block simultaneously. After all blocks have been updated, they are shifted by one-half the width of a block for the next round of updates. This allows two correlated sources that might happen to be located on opposite sides of a border between two neighboring blocks to be updated together on every other update. This blocking provides a means to handle highly correlated searchers/sources, and yields a speed enhancement which is covered in II E.

This simple extension to an MCMC algorithm allows for quick and robust searching of isolated data snippets with up to ~ 100 templates (the limit on the number of templates is due to the computational cost of the multisource F -statistic). In order to be able to handle the entire LISA band, we have to break the search up into subregions containing a manageable number of sources. This introduces the problem of edge effects from sources that lie just outside the chosen search region, but deposit power into the search region. To combat the edge effects we introduced “wings” and an “acceptance window.” The purpose of the wings is to create a buffer between the search region and the regions beyond, so that sources in those outer regions do not adversely affect the searchers in the acceptance window. Searchers that end up in the wings are discarded (even though many might be perfectly good fits to actual sources). This allows us to step through frequency space, using multiple search regions.

A problem, which we call “slamming,” can occur when a bright source lies just outside the search region. Slamming is the tendency of the searchers to migrate to the edge of the search region in an effort to fit the power that bleeds into the search region. Since the search region contains only a portion of the signal from the exterior

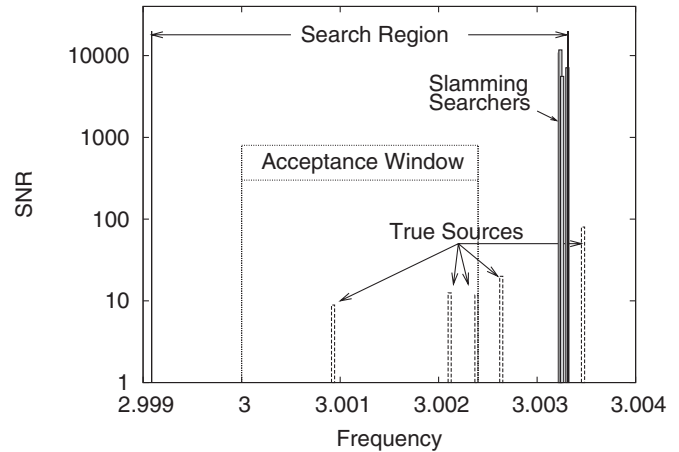


FIG. 7. An example of slamming, in which four searchers are drawn to the edge of the search region by a bright source placed just outside the search region.

source, a single searcher is a poor match, and so other searchers soon are recruited to improve the fit. Figure 7 shows a case of slamming, where four searchers (in a data snippet with four sources) are drawn off to the edge of the search region by a bright source just outside the search region. A tell-tale feature of slamming is the large amplitudes of the searchers and their high degree of correlation or anticorrelation.

One fix to the slamming problem is to weight the matches in the wings less than matches in the acceptance window. We do this by attenuating the contribution from the wings in the noise-weighted inner product. The attenuation is done by increasing the noise spectral density in the wings, which we call “wing noise.” We exponentially increase the noise spectral density, starting at the edge of the acceptance window, as shown in Fig. 6. Figure 8 shows a search using wing noise in the same region as Fig. 7. As can be seen there is no slamming, and the searchers were able to find the true sources.

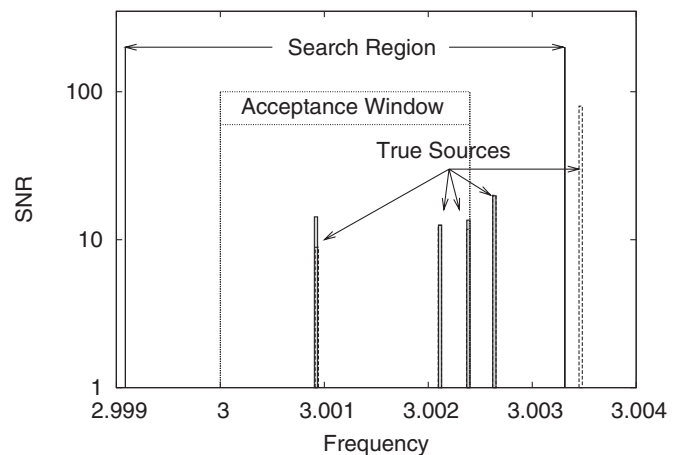


FIG. 8. The same search shown in Fig. 7, but with the inclusion of wing noise. The wing noise has cured the slamming.

Another fix to the slamming problem is to adopt a hierarchical approach that starts by identifying and removing the brightest sources in the data stream. Computationally inexpensive pilot searches can be performed using a small number of searchers, which naturally tend to lock on to the brightest sources. Information gleaned from the pilot searches can then be used to subtract bright sources that might lead to slamming. This hierarchical approach will be described in more detail in the Sec. III A. In general use of our BAM algorithm, we use both wing noise and information about bright sources outside the search regions to lessen the chance of slamming.

E. Time scaling

Since our ultimate goal is to simultaneously solve for $15\,000^+$ galactic binary systems, careful attention has to be given to the computational cost of the algorithm. Even though the BAM algorithm exploits the local nature of the correlations between sources, there can be dozens of highly correlated sources in the $\sim 1\ \mu\text{Hz}$ bandwidth of a single source. Thus, we have to pay particular attention to how the computation cost scales with the number of searchers in a search region.

A straightforward BAM algorithm that does not employ the F -statistic to compute the likelihood or multisource Fisher matrices in the proposal distributions would have a computational cost that scales linearly with the number of sources being searched for, as the computational cost per source would be constant. The goal of a purely linear scaling is unrealizable in our current version of the BAM algorithm for two reasons.

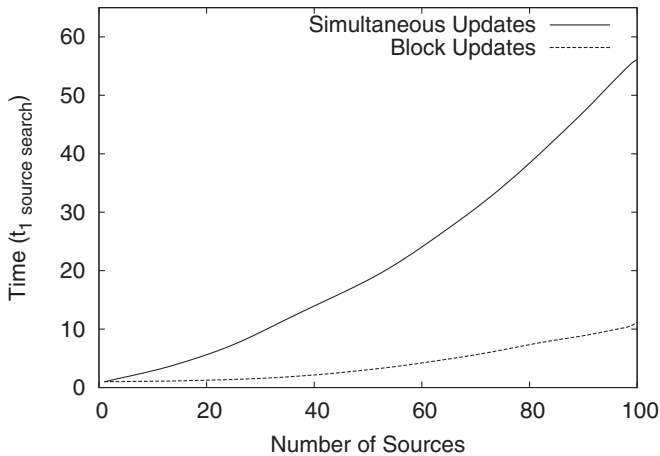


FIG. 9. Plot of the average computational time per update per search template for two cases, one where all sources are updated simultaneously, and the other where block updates are used. The computational time (y-axis) has been scaled by the average time per update for a single source search. The increase in cost per template update grows much more slowly for the blocked search. To achieve linear scaling in the total cost of the BAM algorithm, we would need a constant computational cost per individual template update.

First, in using the F -statistic to calculate the likelihood values we are required to solve Eq. (14) for the time-independent amplitudes a_i . Inverting this equation would scale as N^3 , and while solving for the amplitudes using an iterative method is significantly better it does not scale linearly with N . In future versions of the algorithm, we are considering the option of limiting the usage of the F -statistic to each frequency block, thereby significantly reducing the scale of the linear algebra problem.

The second impediment is that some proposal distributions, such as a multivariate normal distribution jumping along the eigendirections of the variance-covariance matrix, are inherently nonlinear in computational cost (calculation of the Fisher information matrix alone scales as N^2). It is here that the use of the blocks shows great benefit. Figure 9 shows how the time to perform a searcher update is affected by the number of searchers. The plot shows the average cost per individual searcher that is updated in a search using only multivariate normal proposal distributions. In one case, all searchers are updated simultaneously (i.e. the entire search region is treated as one block). In the other case, breaking the updates into the blocks lessens the number of searchers being updated simultaneously, greatly reducing the overall computational cost as the number of sources grows. The search was performed on a one-year data stream in a 0.01 mHz snippet starting at frequency of 3 mHz with blocks that were 4/year wide. Up through $N = 20$, no blocks contained more than one search template and the increase in computational cost per search template update was just 1.25 times greater than a single template search, due solely to the use of the multisource F -statistic. Beyond $N = 20$, the blocks began to contain more than one searcher, and the nonlinear nature of the normal proposal started to drive up the cost per update.

III. ISSUES WHEN USING THE ALGORITHM

Thus far, we have described the features of the BAM algorithm, and discussed how those features aid in optimizing the algorithm. In this section some of the implementation issues will be addressed. We will discuss choices of the sizes of search regions, the sizes of the wings, hierarchical searching, stopping criteria, and false-positive/false-negative levels.

A. Hierarchical method

We are able to perform searches in a hierarchical manner simply by choosing to use fewer search templates than there are sources. For example, if we search a region that contains 10 sources with a multitemplate built from 2 galactic binary templates, the algorithm will generally recover the 2 sources that return the highest likelihood value (usually they will correspond to the highest SNR sources in the region). In this hierarchical approach, the solution for the brightest sources will be thrown off by the sources that were neglected, but this bias can be corrected

at subsequent stages in the hierarchy. At the next stage in the hierarchy more galactic binary templates are used, with the information from the earlier searches providing starting locations for some of the search templates. In practice, we will not know the true number of sources in any given search region; here our estimates of the local source density will inform our choice of a starting number of searchers.

Figure 10 illustrates an example where a $1000f_m$ data snippet containing 281 sources with an observation time of three years was searched using 10, 30, and 50 searchers. In the case using 10 searchers, the 10 sources with the highest SNRs were found, while for the case with 30 searchers only 25 of the 30 highest SNR sources were among those found (the other 5 searchers did find true sources, albeit with lower SNRs).

One may be tempted to subtract the signals of sources found in the pilot search, before looking for more sources in the region, but our experience leads us to believe that this is not the best way to proceed. In a hierarchical search, the initial solution for the bright sources will be thrown off by dimmer sources that they overlap with. To stop these errors from propagating it is better to allow the algorithm to refine the fit to the bright sources at subsequent stages in the hierarchy.

Instead of removing the signal from the data stream, we include the information in the multitemplate represented by the combination of the search templates. Within the framework of our algorithm, one has several options of how to include the information at subsequent stages. One option is to begin the next search with some of the search

templates at the source locations determined at the previous iteration, then allowing them to behave like any other searchers from that point on. Another option is to hold the intrinsic parameters of the previously determined signals fixed during the high heat portion of the annealing phase (the extrinsic parameters of all templates are still updated since we are using a multitemplate F -statistic). The advantage of the latter approach is that it protects the information gleaned from earlier stages in the hierarchy, as there is a danger that a previously determined template will be dislodged during the annealing phase. When the temperature reaches a selected value, the constraints are withdrawn and the inherited search templates get treated like any other. This allows the information gathered in the earlier runs to be preserved, and also allows for adjustments to be made to the parameters as needed due to the inclusion of other searchers.

Another benefit is that such hierarchical searching can be used to improve efficiency. Knowing where the very bright sources are in the data snippet can be used to lessen the sizes of the wings of the search regions. This will be discussed in more detail in the next subsection.

B. Choosing window sizes

There are several factors that influence the choice of window size. In Sec. II E we saw that the cost per template update grows with the number of templates in a window. This argues for using small search windows. However, in Sec. II D we saw that it was necessary to include a wing region to mitigate edge effects, and any signals recovered in the wings had to be discarded. Thus, we would like for

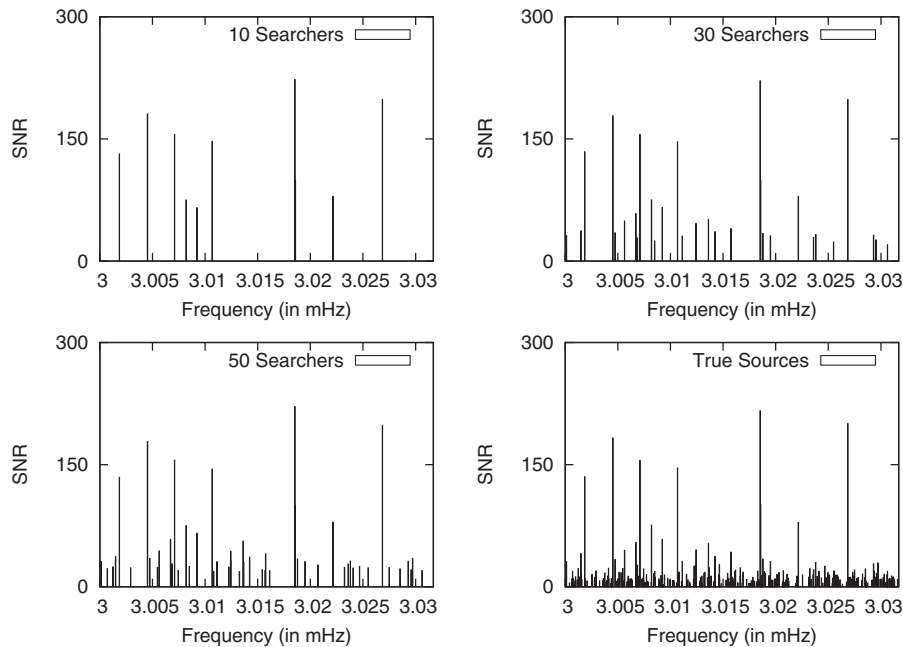


FIG. 10. Plot of the recovered (and true) parameter frequencies and their respective SNRs for the sources recovered by the BAM algorithm searching a $1000f_m$ data snippet containing 281 sources with a 3 yr time of observation using 10, 30, and 50 searchers.

the wings to cover a small fraction of the search window, so that the fraction of templates assigned to wings is small. The size of the wings is determined by the typical bandwidth of a source. A natural choice is to set the wing size equal to the typical half-bandwidth as this ensures that little power from sources located outside of the search region leaks into the acceptance window. The source bandwidth is determined by two factors, the width of the sideband comb,

$$B_c \approx 2 \left(5 + 2\pi f \frac{1\text{AU}}{c} \right) f_m, \quad (17)$$

and the degree of spectral leakage caused by the finite observation time. The spectral leakage depends on the amount the carrier frequency differs from being an integer multiple of $1/T_{\text{obs}}$, and falls off inversely with the number of frequency bins. For observation times on the order of years and for frequencies in the milli-Hertz range, the two contributions to the bandwidth are of similar size. Somewhat larger wings are needed if very bright sources have not been regressed from neighboring frequency windows.

These physical considerations dictate a total wing size of $\sim 15 \rightarrow 30 f_m$, and we would like to use snippets considerably larger than this in order to maximize the fraction of the templates that appear in the acceptance window. The problem with this is that in regions of high source density we end up with template densities of 1 per ~ 4 frequency bins, so for a 1 yr data stream the wings alone account for $4 \rightarrow 7$ templates. Thus, to keep us in the regime where the cost per update is close to that of a single source, we were forced to use snippets where the acceptance window was comparable in size to the wings. In future upgrades we hope to improve the scaling of the algorithm so that we can use larger search windows.

C. Stopping criteria

In LISA data analysis we will not only have to determine source parameters, but also the number of sources that can be resolved. Studies of the galactic confusion noise levels [4–7,46,47] provide some answers concerning source populations across the LISA band, as well as estimates of the number of resolvable binaries as a function of frequency [7]. However, this information is better suited to determining the number of source templates to start with, than the

number with which to end. In order to discover where we must stop we will have to listen to the data.

Models with more parameters will always produce better fits to the data, but beyond a certain point the recovered parameters become meaningless. What we seek is the best fit to the data with the simplest possible model. For a model with uniform priors we seek to minimize

$$\chi^2 = (s - h|s - h), \quad (18)$$

using the smallest number of source templates. Using $\chi^2 = (\Delta h + n|\Delta h + n)$ and $\Delta h = -h_i \Delta \lambda^i$, along with the maximum likelihood estimate for the parameter errors, $\Delta \lambda^i = (\Gamma^{-1})^{ij}(h_j|n)$, where $\Gamma_{ij} = (h_i|h_j)$ is the Fisher information matrix, we find that the expectation value of χ^2 is given by

$$\langle \chi^2 \rangle = \langle (n|n) \rangle - (\Gamma^{-1})^{ij} \Gamma_{ij} = \mathcal{N} - D, \quad (19)$$

where \mathcal{N} is the total number of data points and D is the model dimension. As one might have anticipated, it does not make sense to use a model with more parameters than there are data points. This sets an upper limit of 1 source template every ~ 2 frequency bins, as there are 4 data points per frequency bin (2 independent data channels, each with a real and imaginary part), and 7 parameters per template. More refined criteria, such as the Bayesian evidence, set more stringent stopping criteria.

There have been many different suggestions of how to weight goodness of fit against model complexity. Two in common use are the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) [48]. We tried both, and found neither to be particularly satisfactory as a stopping criteria, as they generally overestimated the number of sources in the data. Instead we settled on the Laplace approximation to the full Bayesian evidence. The evidence $p_X(s)$ for a model X given data s is given by the integral

$$p_X(s) = \int p(s|\vec{\lambda}, X) p(\vec{\lambda}, X) d\vec{\lambda}. \quad (20)$$

Computing this integral is extremely expensive for high dimension models, but the Laplace approximation provides the estimate:

$$p_X(s) \approx p(s|\vec{\lambda}_{\text{MAP}}, X) \left(\frac{\Delta V_X}{V_X} \right), \quad (21)$$

where $p(s|\vec{\lambda}_{\text{MAP}}, X)$ is the maximum of the posterior, V_X is

TABLE I. Source parameter for the 4 sources in the data snippet.

SNR	A (10^{-24})	f (mHz)	θ	ϕ	ψ	ι	φ_0
1.23	2.05	3.001 514 406	1.259	4.012	0.759	1.183	2.551
7.67	10.8	3.000 315 843	2.437	2.753	2.484	2.173	2.550
9.65	12.1	3.000 454 748	2.198	0.422	2.880	2.263	2.991
10.76	9.27	3.001 985 584	1.336	5.863	0.931	2.805	4.048

the volume of the model's parameter space, and ΔV_X is the volume of the uncertainty ellipsoid, which can be estimated from the curvature of the posterior. (To simplify the calculation, we ignored the gradient of the priors and used the maximum likelihood estimate for $\vec{\lambda}$ and the Fisher information matrix as a proxy for the curvature of the posterior.) In general, adding another source template to the model will increase the likelihood, however, the $\Delta V_X/V_X$ term penalizes larger models and serves as a built-in Occam factor.

As an example we will look at a data snippet containing 4 sources. The source parameters and SNRs are shown in Table I. There is one dim source which is not expected to be recovered (see Sec. III D for more details). Figure 11 plots the logarithms of the evidence and likelihood for models of increasing size. One can see that the evidence is peaked at the model with 3 source templates, while the likelihood continues to climb as the model dimension is increased. This tells us that the data favors a model with 3 sources over one with 4.

An alternative approach to model selection is to use the reverse-jump MCMC algorithm [49], which allows for transitions between models of different dimension. The fraction of the time the chain spends exploring each model is used as a measure of the relative evidence for the different models. We plan to use the RJMCMC approach in future versions of the algorithm.

One limitation of the way we have formulated our Bayesian model selection is that the evidence for a zero source model is ill defined, so we cannot compare models with 0 and 1 source templates. This limitation can be removed if we expand our models to include the instrument noise parameters, which we plan to do in future versions of the algorithm.

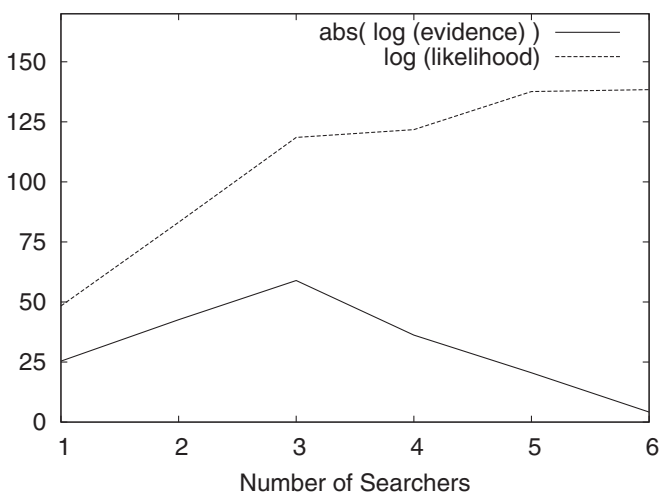


FIG. 11. Plot of the magnitude of the logarithms of the likelihood and evidence for increasing numbers of searchers, searching a data snippet with sources whose parameters are listed in Table I injected into it.

For those that favor the frequentist approach to data analysis, we provide some rough estimates of the false alarm and false dismissal rates in the following subsections. To make such estimates we need to adopt some criteria to decide if a signal has been detected. Since we know the parameters of the injected sources, we based the detection criteria on how well the source parameters were recovered. We found that setting a $5\text{-}\sigma$ error threshold for each parameter (as determined from the Fisher matrix) was a reliable indicator. Setting the threshold anywhere between $4 \rightarrow 10\sigma$ gave almost identical results as the misses tend to be very far from the mark, while the detections have errors compatible with Fisher matrix predictions.

D. False negatives

Here we study the detection rate for the BAM algorithm as a function of the SNR of an isolated source. To this end we performed searches of data snippets between 3 mHz and 3.031 688 mHz (a $1000f_m$ segment of the LISA band), each containing a single source. A set of parameters for 100 sources were chosen at random. Keeping all other parameters constant, the amplitudes were varied to increase the SNR. Simulated data sets for each of the 100 sources were created at different SNR levels. The BAM algorithm was then applied to the simulated data using short chains of 15 000 steps (10 000 steps in the annealing phase and 5000 set in the sampling phase of each run). The searches were also run using longer chains, consisting of 75 000 steps (50 000 annealing/25 000 sampling). For each of two types of chains, two analysis methods were used to determine the parameter values. In one method the parameter values were determined by using average values of the parameters in the chain from the sampling phase (i.e. Bayes estimates). In the second method the parameter values were determined from the mode of the parameter histograms from the sampling phase. Figure 12 shows the detection probabilities for these searches based on the two

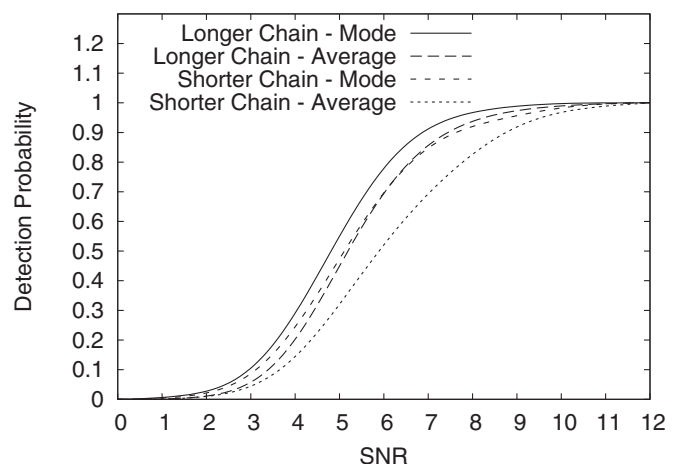


FIG. 12. Probability of detection of a source using the BAM algorithm as a function of source SNR.

analysis methods. The resulting parameter set is called a detection when they deviate by less than $5\text{-}\sigma$ in each of the true source's intrinsic parameters.

As expected, the detection probability depends on the length of the chain, with the longer chains giving higher detection rates. However, a single very long chain is not the best way to ensure detection. Consider the search for sources with $\text{SNR} = 5$ using the parameter mode estimates. The detection rate for the shorter chains was 0.45, while for the longer chain the rate was 0.57. However, if the search for a single $\text{SNR} = 5$ source is repeated multiple times, the detection rate also comes out at ~ 0.45 . Thus, if the short chain search is run twice and the results from the two runs are combined, the detection probability improves to $\sim 1 - 0.55^2 = 0.7$, at a total cost of 30 000 steps, which is still less than half the cost of the long chain searches.

Of the two methods for determining the source parameters, the mode performed better than the Bayes estimate in the low SNR region. On the other hand, the mode is harder to compute, especially when there are a large number of sources, so the current implementation of the full scale BAM algorithm still uses Bayes estimates. This will be corrected in future versions.

E. False positives

We now turn our attention to computing the false alarm rate by searching data streams that contain only instrument noise. We employed two approaches: in the first we perform multiple searches with differing noise realizations and in the second we performed an extended search of one noise realization.

For the first test we performed 20 000 searches using a single search template in the same frequency band we used to study the false negatives. Figure 13 shows a histogram of the SNRs for the finishing points of the searches. This can be used to give an idea as to the level where false positive

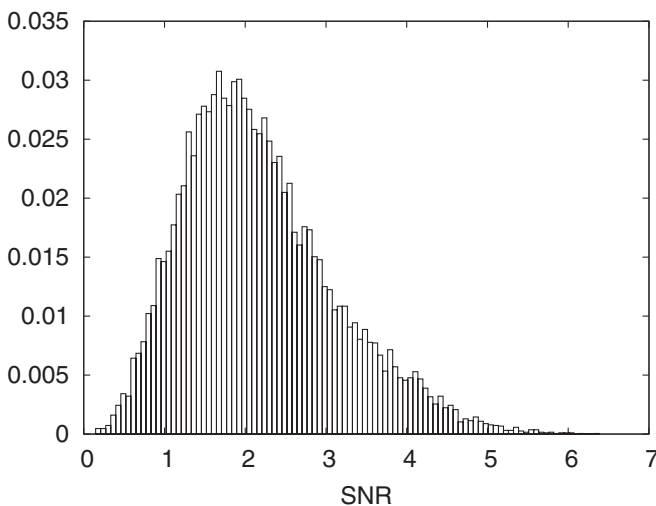


FIG. 13. Histogram of the false-positive SNRs for searches on source-free noise in a $100f_m$ band starting at 3 mHz.

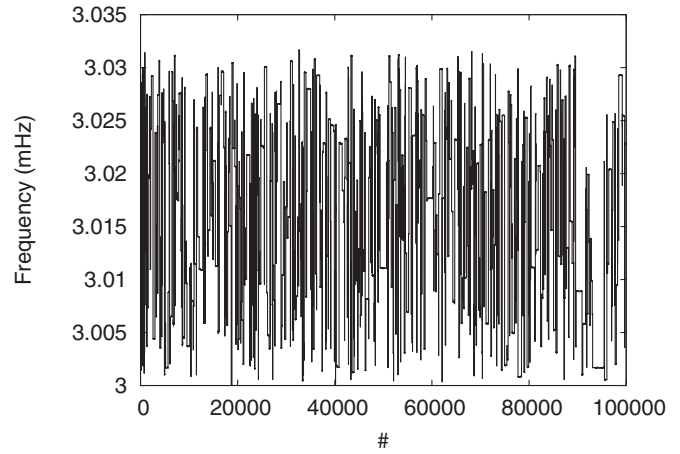


FIG. 14. Plot of the frequency steps in a chain searching source-free noise in a $1000f_m$ band starting at 3 mHz.

begins to become a concern. In this frequency range, the histogram tells us that more than 99% of the searches ended on parameters leading to a SNR less than 5. So in accepting any result from a search with SNR above 5 in this regime there is roughly a 1% chance of accepting a false positive, with the probability dropping precipitously for searchers returning higher SNRs. We repeated the analysis at different frequencies and found the false-positive level to be much the same across the LISA frequency band.

The second test used the same $1000f_m$ band at 3 mHz, but now a single long chain of 1×10^6 steps was performed. Figure 14 shows the frequency parameter over 100 000 steps in the chain. One notices that the chain does not lock in on any particular frequency, rather it continually wanders about the entire search region. While this is common, it is not always the case, as instrument noise can sometimes mimic a monochromatic source

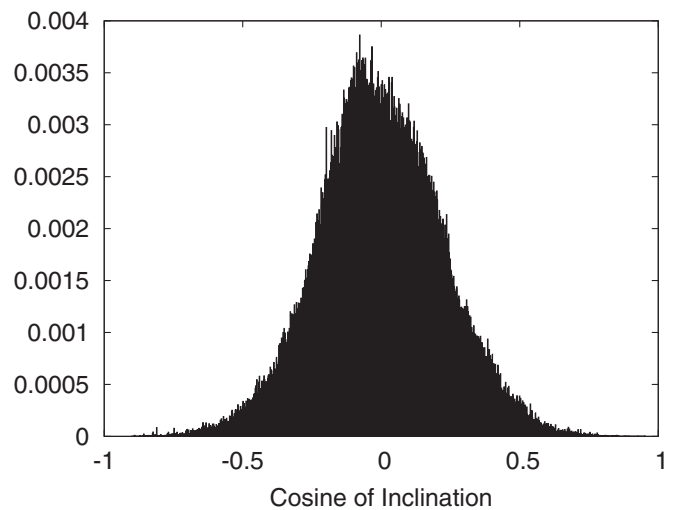


FIG. 15. Histogram of the cosine of the inclination angle for a chain searching source-free noise in a $1000f_m$ band starting at 3 mHz.

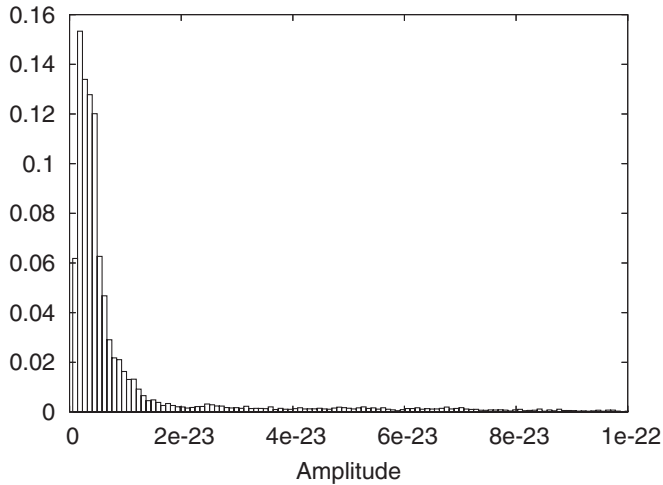


FIG. 16. Histogram of the amplitude parameter for a chain searching source-free noise in a $1000f_m$ band starting at 3 mHz.

well enough to slow or even stop such exploratory movement in a chain. Two other reliable indicators of a false positive are the amplitude and the inclination angle of the binary system. Figures 15 and 16 show their respective histograms. The cosine of the inclination angle is peaked around zero, and the amplitude is peaked at a level that gives templates whose spectral density has the same magnitude as the noise. The reason for these preferences is that they allow the template to optimally match the noise in the two LISA data channels. An inclination angle of $\pi/2$ gives equal weight to the two channels, which allows the template to match the noise level in both channels equally well.

To summarize, if the BAM algorithm recovers a source with $\text{SNR} \geq 5$ it is unlikely to be a false positive. Moreover, the chains show some very characteristic features when the templates are just matching noise, and these features can be used as a diagnostic to exclude false positives. Lastly, one can always run the search multiple times, and if the same set of parameters are recovered over and over again it is less likely that we have found a false positive.

IV. EXAMPLE SEARCHES

In this section we show the results of several types of searches performed with the BAM algorithm. While results for single source searches are easy to describe, when the search is for hundreds of sources, and there are hundreds of successes, the numbers are much harder to present in a condensed manner. To that end, we will present multiple search cases using plots like those in Fig. 10, where the frequency values are shown, and we will list the deviations of the recovered intrinsic parameters scaled by the uncertainties given by the Fisher information matrix for the sources. Since we are mostly interested in the search phase of the algorithm the post-search MCMC runs we chose to be rather short, so the recovered posteriors are a little

ragged. With this in mind we set our cutoff criteria for “finding” a source at a deviation of $5\text{-}\sigma$ from any one of the true source’s intrinsic parameters. In practice this cutoff was fairly robust as any template that did not have all parameters within a few σ of the source typically had one or more parameters that were tens to hundreds of σ out. We also imposed a SNR minimum of 5 in addition to the Bayesian evidence criteria. This was perhaps stricter than necessary to keep down the possibility of a false positive (and in fact did lead to one actual detection being discounted), but with that cutoff there were no instances of false positives in any of the searches presented here.

A. Searches of the mock LISA data challenge training sets

The mock LISA data challenge (MLDC) consists of several types of challenges for the LISA data analysis community to test search algorithms using simulated LISA data. The first round in the MLDC consists of challenges for three source types: galactic binaries, supermassive black holes, and extreme mass ratio inspirals. For this work we will focus on the challenges dealing with galactic binaries. Provided with each challenge are two data sets, one is a blind test where the source parameters are unknown, while the other is an open test where the source parameters are provided so that one may synchronize conventions. In this paper we will only be discussing results from the open data sets, as the MLDC task force has asked that all results for the blind challenges be embargoed until December 2006.

1. Single binary searches

The first challenges consist of single binary systems injected into a LISA data stream with instrumental noise. In Challenge 1.1.1 there are three tests. The first is a monochromatic binary with frequency, $f = 1.0 \pm 0.1$ mHz, the second has a frequency of $f = 3.0 \pm 0.1$ mHz, and the third a frequency of $f = 10.0 \pm 1.0$ mHz. While the BAM algorithm is designed to handle multiple source searches, these single source challenges provide a means to test our conventions as well as check for any modeling errors. One source of modeling error is that we used a low frequency detector response model, which restricts us to searching for signals below 7 mHz. Thus we have only performed searches on the first two of these single source challenges.

In the 1 mHz case, the true source parameters and the results from the BAM algorithm are given in Table II. For this work all uncertainties will be calculated using a Fisher information matrix at the parameter values given by the chain. A longer run of the data chains after burn-in could also provide a means to calculate these uncertainties, but as was shown in [17] for the intrinsic parameters these will be a very good match to those from the Fisher calculation. Here the deviations (or discrepancies) between the true

TABLE II. Results of a search of the MLDC training data set 1.1.1a.

	A (10^{-22})	f (mHz)	θ	ϕ	ψ	ι	φ_0
True parameters	1.789 229 908	0.993 034 853 5	0.474 114 326 8	5.199 21	3.975 816	0.179 395 6	5.781 211
Recovered parameters	2.364	0.993 034 139	0.4575	5.196	4.475	0.7364	4.832
Parameter uncertainties	0.3876	9.493×10^{-7}	0.029 83	0.017 54	0.3126	0.2095	0.6234

parameter values of the three intrinsic parameters and the values determined through the search are less than 1σ : $\Delta f \equiv (f_{\text{true}} - f_{\text{search}}) = -0.7530\sigma_f$, $\Delta\theta = -0.5563\sigma_\theta$, and $\Delta\phi = -0.20498\sigma_\phi$.

For the 3 mHz case, the search returned mean parameter values that were also discrepant from the true source parameters by less than 1σ ($\Delta f = 0.2164\sigma_f$, $\Delta\theta = -0.09466\sigma_\theta$, $\Delta\phi = -0.7860\sigma_\phi$).

In these two searches, the algorithm is performed exactly as expected. This suggests that our current implementation of the algorithm is free of any significant systematic errors in either waveform generation or calculation of the likelihood values, as the MLDC data was created with an independently developed code.

2. Low source confusion

Challenge 1.1.4 for the MLDC is a test for algorithms in the low source confusion regime, where the source density is ~ 1 source per $10f_m$. Results for a search of the training data set are presented here.

Using only the frequency range given in the challenge ($3.000 \text{ mHz} < f < 3.015 \text{ mHz}$), the BAM was run on the data stream (an approximate range of the number of source was given in the challenge, and indeed the exact number in the training data is known to be 50, but that information was not used in directing the algorithm). The log evidence was used as the stopping criteria for each search region. The frequency range of the data snippet was divided into 20 search regions such that each acceptance window was $0.75 \mu\text{Hz}$ in width, and the width of the wings were 0.1 times the bandwidth of a typical source with a frequency in the search region.

Figure 17 gives a visual representation of the results of the search. The upper left plot gives the locations in frequency of the 50 individual sources in the data snippet. The heights of the bars show the SNR of each source, while the width of the bars is $1f_m$. Results of the search are similarly expressed in the upper right plot. As these two plots are very similar, the lower left plot has been provided to highlight the difference between them. The lower right plot shows how well the search algorithm fit the true source parameters. It is a histogram of the differences in

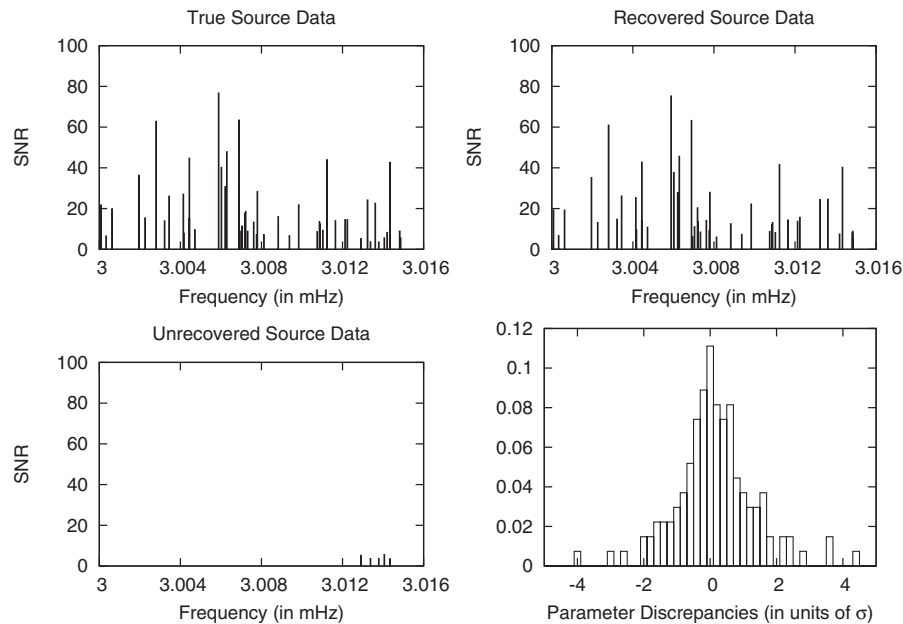


FIG. 17. Plots detailing the results of a search of the LISA data stream for training data set 1.1.4 of the mock LISA data challenge. The upper two plots show the frequencies and SNRs for the true and recovered sources. The lower left plot gives the frequencies and SNRs of the source not recovered by the BAM algorithm. The lower right plot shows the discrepancies between the true and recovered intrinsic parameters in units of the parameter variances.

the 135 intrinsic parameters of the recovered sources in units of their respective variances (as calculated via a Fisher information matrix located at the recovered parameter values). Nearly 92% of the parameters recovered by the algorithm differed from their true values by less than 2σ . This fit might be improved some with a so-called “finisher” step, which will be briefly discussed in more detail in the following subsection.

The five unrecovered sources shown in the lower left plot of Fig. 17 represent the false negatives of this particular search. Three of these sources had SNRs < 5 and thus were not expected to be recovered given the results of Sec. III D. The remaining two unrecovered sources had SNRs < 5.6 .

3. Strong source confusion

Challenge 1.1.5 for the MLDC is a test for algorithms in the high source confusion regime, where the source density is ~ 1 source per $\sim 2.5f_m$. Results for a search of the training data set are presented here.

The BAM algorithm was run on the data stream using the frequency range given in the challenge ($2.9985 \text{ mHz} < f < 3.0015 \text{ mHz}$). Again, the approximate range of the number of sources that was given in the challenge was not used in directing the algorithm. The log evidence was used as the stopping criteria for each search region. The frequency range of the data snippet was divided into 10 search regions such that each acceptance window was $0.3 \mu\text{Hz}$ in width, and the width of the wings were 0.5 times the bandwidth of a typical source in the search

region, $\text{BW} = 28f_m$. The wing size of this search is considerably larger than was used in the previous example. Initial runs on this data set with smaller wings showed signs of slamming, so the wing size was increased.

Figure 18 gives a visual representation of the results of the search. The upper left plot gives the locations in frequency of the 44 individual sources in the data snippet. The heights of the bars show the SNR of each source, while the width of the bars is $1f_m$. The density of sources is even higher than first appears in the plot, however, since four sources share an identical frequency ($2.998\,999\,384 \text{ mHz}$) as do three other pairs of sources ($3.000\,085\,802$, $3.000\,629\,082$, and $3.001\,173\,008 \text{ mHz}$).

Results of the search are expressed in the upper right plot of Fig. 18. As can be seen in the plot, the extremely high density of sources prevents the algorithm from recovering all of the sources. Of the 44 sources injected into the data stream, 27 were recovered, which corresponds to a recovered source density of 1 per 4 frequency bins.

The lower left plot shows the sources that were not recovered in the search. Of the 17 unrecovered sources, 13 had a nearest neighbor that was within $1f_m$, including 5 of the unrecovered sources that shared an identical frequency with at least one other source. Sources that are close in frequency are much more likely to be highly correlated than those that are well separated (e.g. the brightest of the sources at $f = 2.998\,999\,384 \text{ mHz}$ anti-correlates with two of the other sources at that frequency at the level of -0.81 and -0.67). This high density and corresponding high levels of correlation introduces two

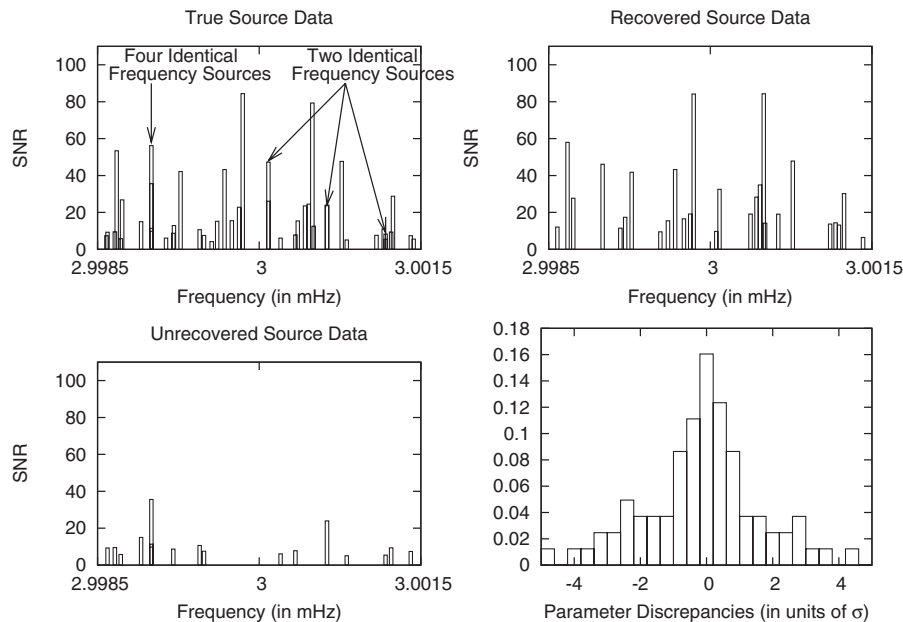


FIG. 18. Plots detailing the results of a search of the LISA data stream for training data set 1.1.5 of the mock LISA data challenge. The upper two plots show the frequencies and SNRs for the true and recovered sources. The lower left plot gives the frequencies and SNRs of the source not recovered by the BAM algorithm. The lower right plot shows the discrepancies between the true and recovered intrinsic parameters in units of the parameter variances.

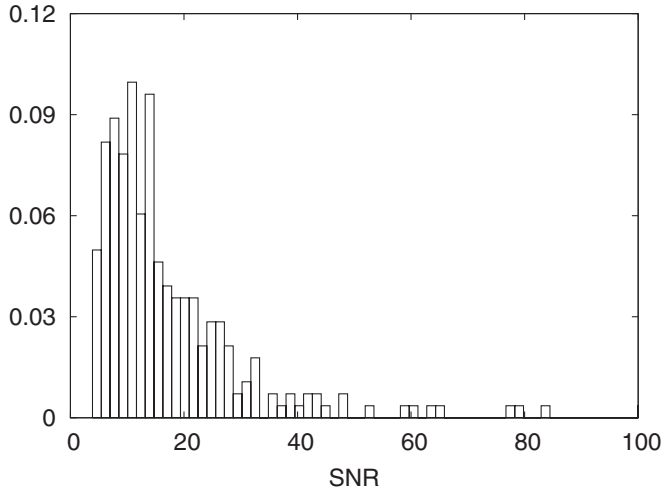


FIG. 19. Partial histogram of the SNRs for the true source parameters of a $1000f_m$ data snippet containing 281 sources with a 3 yr time of observation.

problems for the current implementation of the BAM algorithm. First is that analysis of the chains was done using the mean values of a particular search template. With nearby sources, the individual searchers can jump between the sources and the calculated mean value is a weighted mean of the two, or more, close sources (weighted by the number of steps in the chain spent at each source). In the next implementation of the algorithm, we combine all the parameter chains of a given type into a single histogram

and use standard spectral line fitting software to fit the combined posterior distribution function by multiple Gaussians. Second, the current implementation of the algorithm includes a “blast” proposal distribution that separates highly anticorrelated sources ($\kappa < -0.9$). This proposal was included to lessen the effect of slamming (by performing a uniform draw on one of the two anticorrelated searchers). With the inclusion of the wing noise and returning to a 7 parameter search (per template), this proposal will most likely not be needed. This should allow the search templates to spend more time in areas with highly anticorrelated sources.

The lower right plot shows how well the search algorithm fit the true source parameters. It is a histogram of the differences in the 81 intrinsic parameters of the recovered sources in units of their respective variances (as calculated via a Fisher information matrix located at the recovered parameter values). The spread of the discrepancies for the recovered sources is larger than those for the case of low source confusion shown in Fig. 17. Just over 76% of the parameters recovered by the algorithm differed from their true values by less than 2σ . This departure from the Fisher matrix predictions is due to the additional confusion noise from unrecovered sources.

Lastly, the fit could be improved using a finisher step in the analysis process. While the full details of such a finisher are beyond the scope of this paper, one can think of it as continuing the search algorithm using proposals specific to providing efficient mixing of the chain (such as

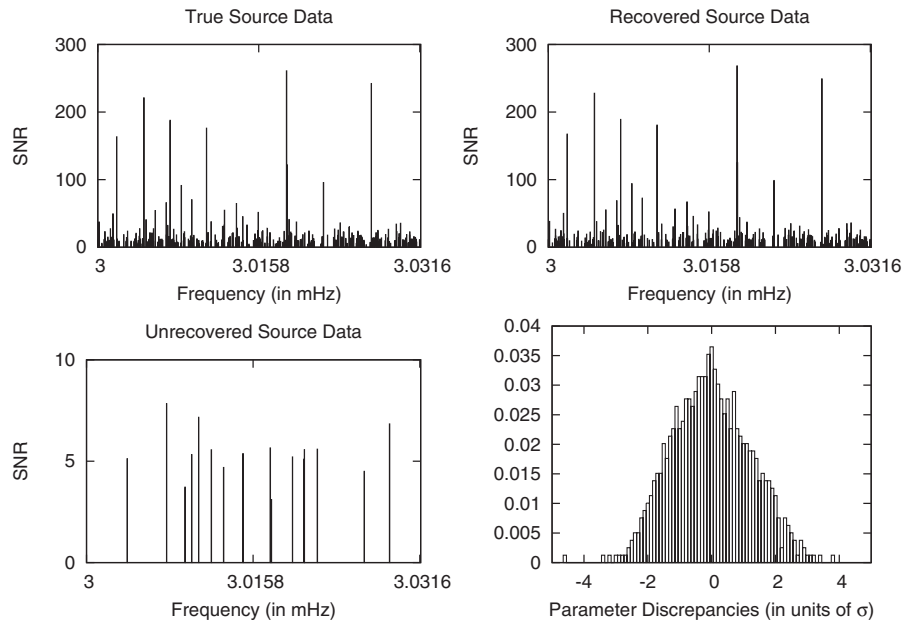


FIG. 20. Plots detailing the results of a search of the LISA data stream with a 3 yr time of observation containing 281 sources drawn from a galactic distribution. The upper two plots show the frequencies and SNRs for the true and recovered sources. The lower left plot gives the frequencies and SNRs of the source not recovered by the BAM algorithm. The lower right plot shows the discrepancies between the true and recovered intrinsic parameters in units of the parameter variances.

a drawing from a multivariate Gaussian distribution) that will allow for searchers to work through the issues created by the high levels of correlation and reach the posterior distribution. Also, in this step sources can be introduced to the fit given by the BAM algorithm to provide a better fit.

B. A Large N search for resolvable binaries

In this subsection we will discuss the results of a search for sources in a $1000f_m$ data snippet at 3 mHz. These sources were chosen from a galactic model described by Nelemans *et al.* [6]. For this realization there are 281 sources in the data snippet.

Unlike the previous searches, the observation time is 3 years. This provides a test of the algorithm for multiyear observation times, and models a search through a nontrivial section of the galactic background (nearly 1% of the overall frequency range, and $>1\%$ of the expected resolvable sources). With a 3 yr observation time, all but four of the sources have a $\text{SNR} > 5$. Figure 19 shows a histogram of the SNRs for sources below 100 (another 9 sources have $\text{SNRs} > 100$).

The upper left plot of Fig. 20 shows the locations in frequency of the individual sources in the $1000f_m$ data snippet. The heights of the bars show the SNR of each source, while the width of the bars is $1f_m$, while the results of the search are shown in the upper right plot. As these two plots are very similar, the lower left plot is provided (and rescaled) to highlight the difference between them. The search was able to find 365 of the 281 sources, with all but 3 of the unrecovered sources having a $\text{SNR} < 6$. One of these “unrecovered” sources was an instance of the SNR cutoff inadvertently omitting a detected source (the source at $f = 3.02634519$ has a $\text{SNR} = 4.50$, and it was recovered with a $\text{SNR} = 4.62$). It is interesting to note that these results are consistent with the prediction in Sec. III D regarding the rate of false negatives. This suggests that most of the 12 sources that were not recovered and have a $\text{SNR} > 5$ might be found with repeated searches.

The lower right plot of Fig. 20 shows a histogram of the differences in the 795 intrinsic parameters of the recovered sources in units of their respective variances (as calculated via a Fisher information matrix located at the recovered parameter values). Slightly more than 91% of the parameters recovered by the algorithm differed from their true values by less than 2σ .

V. CONCLUSION

We have developed and tested an algorithm that is capable of locating and characterizing galactic binaries across the entire LISA band. In regions of strong source confusion, we found that the algorithm could recover 1 source every 4 frequency bins. We found that the algorithm performs very well on snippets taken from a full galactic foreground model, and since the BAM algorithm develops a global solution by sewing together searches over subsets of the LISA data, completing the full analysis of the galactic simulation is just a matter of computer time.

While the current algorithm is very effective, we identified many improvements and extensions that are now being implemented. The waveform modeling has been updated to include the full LISA response and frequency evolution, and we have reverted to performing full parameter searches due to the computational cost incurred by the multisource F -statistic. Work is currently in progress to extend the search to include parameters that describe the noise in each data channel. This extension will be particularly important below 3 mHz as the effective noise level will be set by unresolved sources, so we will not know in advance what weighting to use in the inner products. Our current stopping criteria using the Laplace approximation to the Bayes evidence could be eliminated if we switch to a transdimensional reverse-jump MCMC method [50]. The analysis of the post-search chains can also be improved using spectral line fitting techniques.

Even with our current algorithm, we estimate that it would take less than two weeks to process a full galactic foreground on a 3 GHz, 128 node cluster. With the modifications we are implementing we expect both the speed and fidelity of the algorithm to be much improved. We will have an opportunity to test the updated BAM algorithm on a full galactic simulation in the second round of mock LISA data challenges, which are set for release in January 2007.

ACKNOWLEDGMENTS

This work was supported at MSU by NASA Grant No. NNG05GI69G. A portion of the research described in this paper was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

-
- [1] P. Bender *et al.*, LISA Pre-Phase A Report, 1998.
 - [2] C.R. Evans, I. Iben, and L. Smarr, *Astrophys. J.* **323**, 129 (1987).
 - [3] V.M. Lipunov, K.A. Postnov, and M.E. Prokhorov, *Astron. Astrophys.* **176**, L1 (1987).

- [4] D. Hils, P.L. Bender, and R.F. Webbink, *Astrophys. J.* **360**, 75 (1990).
- [5] D. Hils and P.L. Bender, *Astrophys. J.* **537**, 334 (2000).
- [6] G. Nelemans, L.R. Yungelson, and S.F. Portegies Zwart, *Astron. Astrophys.* **375**, 890 (2001).

- [7] S. Timpano, L. J. Rubbo, and N. J. Cornish, *Phys. Rev. D* **73**, 122001 (2006).
- [8] A. Vecchio, *Phys. Rev. D* **70**, 042001 (2004).
- [9] R. N. Lang and S. A. Hughes, *Phys. Rev. D* **74**, 122001 (2006).
- [10] E. Berti, A. Buonanno, and C. M. Will, *Classical Quantum Gravity* **22**, S943 (2005).
- [11] K. J. Rhoads and J. S. B. Wyithe, *Mon. Not. R. Astron. Soc.* **361**, 1145 (2005).
- [12] S. M. Koushiappas and A. R. Zentner, *Astrophys. J.* **639**, 7 (2006).
- [13] L. Barack and C. Cutler, *Phys. Rev. D* **69**, 082005 (2004).
- [14] J. R. Gair, L. Barack, T. Creighton, C. Cutler, S. L. Larson, E. S. Phinney, and M. Vallisneri, *Classical Quantum Gravity* **21**, S1595 (2004).
- [15] S. Drasco, *Classical Quantum Gravity* **23**, S769 (2006).
- [16] S. Drasco and S. Hughes, *Phys. Rev. D* **73**, 024027 (2006).
- [17] N. J. Cornish and J. Crowder, *Phys. Rev. D* **72**, 043005 (2005).
- [18] N. J. Cornish and E. K. Porter, *Classical Quantum Gravity* **23**, S761 (2006).
- [19] E. D. L. Wickham, A. Stroeer, and A. Vecchio, *Classical Quantum Gravity* **23**, S819 (2006).
- [20] A. Stroeer, J. Gair, and A. Vecchio, *AIP Conf. Proc.* **873**, 444 (2006).
- [21] N. J. Cornish and E. K. Porter, *Phys. Rev. D* **75**, 021301 (2007).
- [22] J. Crowder, N. J. Cornish, and J. L. Reddinger, *Phys. Rev. D* **73**, 063011 (2006).
- [23] N. J. Cornish and S. L. Larson, *Phys. Rev. D* **67**, 103001 (2003).
- [24] L. J. Rubbo, N. J. Cornish, and R. W. Hellings, *AIP Conf. Proc.* **873**, 489 (2006).
- [25] C. J. Cutler, M. Vallisneri, and D. A. Brown (unpublished).
- [26] S. D. Mohanty and R. K. Nayak, *Phys. Rev. D* **73**, 083006 (2006).
- [27] J. Gair and L. Wen, *Classical Quantum Gravity* **22**, S445 (2005).
- [28] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).
- [29] W. K. Hastings, *Biometrics* **57**, 97 (1970).
- [30] D. Gamerman, *Markov Chain Monte Carlo: Stochastic Simulation of Bayesian Inference* (Chapman & Hall, London, 1997).
- [31] J. Crowder and N. J. Cornish, *Phys. Rev. D* **70**, 082004 (2004).
- [32] N. Christensen and R. Meyer, *Phys. Rev. D* **58**, 082001 (1998).
- [33] N. Christensen and R. Meyer, *Phys. Rev. D* **64**, 022001 (2001); N. Christensen, R. Meyer, and A. Libson, *Classical Quantum Gravity* **21**, 317 (2004); C. Roever, R. Meyer, and N. Christensen, *Classical Quantum Gravity* **23**, 4895 (2006); J. Veitch, R. Umstatter, R. Meyer, N. Christensen, and G. Woan, *Classical Quantum Gravity* **22**, S995 (2005); R. Umstatter, R. Meyer, R. J. Dupuis, J. Veitch, G. Woan, and N. Christensen, *Classical Quantum Gravity* **21**, S1655 (2004).
- [34] N. Christensen, R. J. Dupuis, G. Woan, and R. Meyer, *Phys. Rev. D* **70**, 022001 (2004); R. Umstatter, R. Meyer, R. J. Dupuis, J. Veitch, G. Woan, and N. Christensen, *AIP Conf. Proc.* **735**, 336 (2004).
- [35] C. Andrieu and A. Doucet, *IEEE Transactions on Signal Processing* **47**, 2667 (1999).
- [36] R. Umstatter, N. Christensen, M. Hendry, R. Meyer, V. Simha, J. Veitch, S. Viegand, and G. Woan, *Classical Quantum Gravity* **22**, S901 (2005).
- [37] L. S. Finn, *Phys. Rev. D* **46**, 5236 (1992).
- [38] C. Andrieu, N. De Freitas, A. Doucet, and M. Jordan, *Mach. Learn.* **50**, 5 (2003).
- [39] *Markov Chain Monte Carlo in Practice*, edited by W. R. Gilks, S. Richardson, and D. J. Spiegelhalter (Chapman and Hall, London, 1996).
- [40] P. Jaranowski, A. Krolak, and B. F. Schutz, *Phys. Rev. D* **58**, 063001 (1998).
- [41] C. Cutler, *Phys. Rev. D* **57**, 7089 (1998).
- [42] N. J. Cornish and L. J. Rubbo, *Phys. Rev. D* **67**, 022001 (2003).
- [43] L. J. Rubbo, N. J. Cornish, and O. Poujade, *Phys. Rev. D* **69**, 082003 (2004).
- [44] S. Geman and D. Geman, *IEEE Trans. Pattern Anal. Machine Intelligence* **6**, 721 (1984).
- [45] C. S. Jensen, A. Kong, and U. Kjaerulff, *Intl. J. Human Computer Studies* **42**, 647 (1995).
- [46] A. J. Farmer and E. S. Phinney, *Mon. Not. R. Astron. Soc.* **346**, 1197 (2003).
- [47] L. Barack and C. Cutler, *Phys. Rev. D* **70**, 122002 (2004).
- [48] G. Schwarz, *Ann. Stat.* **5**, 461 (1978).
- [49] K. A. Arnaud, S. Babak, J. G. Baker, M. J. Benacquista, N. J. Cornish, C. Cutler, S. L. Larson, B. S. Sathyaprakash, M. Vallisneri, A. Vecchio, and J.-Y. Vinet, *AIP Conf. Proc.* **873**, 619 (2006).
- [50] P. J. Green and A. Mira, *Biometrika* **88**, 1035 (2001).