Bayesian Distance Prior in LISA Gravitational Wave Data Analysis

Michael Tauraso

A thesis
submitted in partial fulfillment of the
requirements for the degree of

Master of Science

University of Washington

2024

Committee:
Marilena Loverde
Joey Shapiro Key
Masha Baryakhtar
Tyson Littenberg

Program Authorized to Offer Degree:
Department of Physics

University of Washington

Abstract

Bayesian Distance Prior in LISA Gravitational Wave Data Analysis

Michael Tauraso

Chair of the Supervisory Committee:
Marilena Loverde
Department of Physics

The Laser Interferometer Space Antenna (LISA) mission is expected to detect many Ultra-Compact Binaries (UCBs) in the Milky Way galaxy. In order to link these observations to existing models of star formation and evolution, the spatial distribution of UCBs is an important target for LISA data analysis. In this work `GBMCMC`, an existing code that finds UCBs in LISA data, is extended to incorporate a Bayesian prior on the distance to UCBs from Earth using a simple model of the Milky Way. The effectiveness of this approach is analyzed using simulated data from the LISA Data Challenge.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

LDC: The Lisa Data Challenge

LISA: The Laser Interferometer Space Antenna

UCB: Ultra-Compact Binary

GW: Gravitational Wave

`GLASS`: Global LISA Analysis Software Suite, a program which solves the LISA data analysis problem for multiple source types

`GBMCMC`: Galactic Binary Markov Chain Monte-Carlo, a program which solves the LISA data analysis problem for Ultra-Compact Binary sources and noise.

MCMC: Markov Chain Monte Carlo

RJMCMC: Reversible Jump Markov Chain Monte Carlo, a version of MCMC which allows the model being fit to have a variable number of parameters.

Chapter 1

BACKGROUND

The Laser Interferometer Space Antenna (LISA) is a planned gravitational wave observatory slated to launch in the coming decades. Unlike existing ground based observatories, LISA will be most sensitive in the milli-Hertz (mHz) frequency band[2]. Current astrophysics predicts several types of sources which will produce gravitational waves (GWs) in this frequency band, among them ultra-compact binaries, Extreme Mass Ratio In-spirals (EMRIs), and massive black hole binaries.

The LISA data will contain overlapping mHz gravitational wave signals from throughout the universe. This situation presents a data analysis problem, where observational results from LISA are dependent on the capacity to separate out the signals for individual GW sources. The vast majority of these signals overlap in time and frequency space. This problem has been called the "cocktail party problem" in reference to similar difficulty of separating several overlapping conversations in a crowded room[3].

Ultra-compact binary systems (UCBs) are expected to be the most prolific, with potentially millions of objects contributing to the LISA interferometer signal. UCBs are primarily white dwarf pairs; however, those detectable by LISA may be made up of other stellar remnants or even variable stars orbiting closely with a compact companion, such as AM Canum Venaticorm (AMCVn) stars. For LISA, the vast majority of UCB systems are expected to be continuous sources of GWs, with a similar template waveform. Out of the millions contributing to the interferometer signal, approximately $\mathcal{O}(10^4)$ sources are expected to be individually detectable[16].

In addition to detecting UCBs, LISA will be sensitive to several other classes of object,

and be able to perform some consistency checks on general relativity. Because GWs from all sources and source types will be mixed together in the LISA signal, it is critical that UCB signals are well characterized, because they represent a significant confusion background for other science goals. In order to gain this accuracy, it is necessary to fit all sources at once, as well as the detector noise, in the interferometer signal[13].

In order to help validate various software analysis suites, the LISA Consortium has simulated gravitational wave data as part of the Lisa Data Challenge (LDC)[1]. These simulated data contain both the original parameters of the simulated sources, as well as the interferometric signals expected to be received by the spacecraft. These data allow researchers to check the efficacy of any eventual solution to the global fit problem. For this work I exclusively use the simulated data in LDC's "Radler" release[1].

1.1 How `GLASS` *and* `GBMCMC` *work*

The Global Lisa Analysis Software Suite (`GLASS`) is a prototype code that attempts to solve the LISA global fit problem using custom Markov Chain Monte Carlo methods. GLASS operates on essentially a two-level model, with a high-level Markov Chain Monte Carlo (MCMC) sampler which operates several individual lower-level MCMC samplers which correspond to each source type[13].

The top level MCMC uses a blocked sampling scheme where a block of parameters of the fit are allowed to vary in each iteration, while holding the other parameters fixed. The sampler then cycles through which block is allowed to change. Each block is its own MCMC sampler. A typical block will search for a particular source type, perhaps in a particular frequency range. A core idea of the `GLASS` architecture is that the individual block MCMC samplers are essentially independent. They can each use their own data representation of the LISA GW signal, model distinct GW or noise sources, and define their own likelihood functions. These individual blocks can therefore be developed and tested independently.

---

[1] `LDC1-4_GB_v2.hdf5` in particular.

The focus of this work is on extending `GBMCMC`. `GBMCMC` is an individual block sampler which models the GW signals from UCBs in a narrow frequency band and the confusion noise which arises from unresolvable background UCBs in that same frequency band. Running the full `GLASS` suite is computationally intensive, and results presented here are exclusively from `GBMCMC`.

`GBMCMC` when operated separately from `GLASS` produces posteriors for UCB sources and a noise model. Posteriors for the UCB sources and noise model are then post-processed into a catalog of likely sources[14]. Within `GBMCMC`, and in the resulting catalog, UCBs were parameterized by a vector that includes their principle frequency f_0, its rate of change $\dot{f}$, the amplitude of the gravitational wave $\mathcal{A}$, and parameters describing the sky location, orientation, phase, and frequency rates of change. Prior to this work, `GBMCMC` incorporated priors on sky location and signal-to-noise ratio (SNR) in order to perform an MCMC search in this parameter space using a Bayesian target function[14].

1.2 Bayesian Distance Prior

A catalog of UCBs parameterized by amplitude $\mathcal{A}$ is useful for gravitational wave data analysis, and helps with the problem of UCBs being a confusion signal to other sources; however, as a data product intended for astrophysical study, a catalog of UCB sources parameterized by GW amplitude is less useful. UCBs in the Milky Way are the result of stellar formation and evolution processes, as well as galactic dynamical processes. The models we have for their formation and abundance are framed in terms of their physical properties rather than their gravitational wave emissions.

In an idealized non-interacting[2] UCB $\mathcal{A}$ is a function of principle frequency f_0, chirp mass $\mathcal{M}_c$, and luminosity distance D_L. Additionally, the frequency change due to gravitational waves is a function of $\mathcal{M}_c$ and f_0. These relations give us a path for determining physical parameters of an object from the gravitational waves. Attempts have already been made to

[2]The interacting case is covered in chapter 2

marginalize posteriors from GBMCMC over priors in D_L and $\mathcal{M}_c$ in order to produce a catalog with these physical parameters[10, 8].

This approach of post-processing in a physical prior has the weakness that the search has already occurred using an prior on SNR, which functions as an implicit prior on gravitational wave amplitude and other parameters[12]. Though the effects of a well-chosen SNR prior appear to have very little influence on the sources found [14, 12], such a prior is nonphysical and its influence is nontrivial to remove from the posterior samples, if possible at all.

The goal of this work is to perform a more physically motivated MCMC search in GBMCMC by incorporating prior information on D_L from a model of the galaxy, and building the capability to add prior information on $\mathcal{M}_c$ and $\dot{f}$ to GBMCMC in future work. In order to incorporate this prior information, GBMCMC has been modified to search in a different parameter space, described in chapter 2. Chapter 3 reviews how detailed balance is achieved in the trans-dimensional MCMC algorithm in use by GBMCMC. Modifications to proposals and priors are described in chapters 4 and 5 respectively. Chapter 6 presents the results of these modifications, and chapter 7 concludes with a discussion of possible future work.

Chapter 2

PARAMETERS

In `GBMCMC` gravitational waves from ultra-compact binaries are parameterized by the following parameters:

- $f_0, \dot{f}$ The frequency of the gravitational wave at start of observations, and its time derivative.
- $\mathcal{A}$ The amplitude of the gravitational wave at the detector
- θ, ϕ The latitude and longitude angle in ecliptic coordinates of the source.
- ι, ψ Inclination and polarization of the source UCB to the ecliptic plane.
- ϕ_0 Phase of the gravitational wave at start of observations.

Physically, UCBs have a chirp mass $\mathcal{M}_c$ and a luminosity distance D_L, which together determine the relationship between the underlying engine of gravitational wave emission and the gravitational waves received at earth. This relationship is very similar to luminosity/distance relationships used in optical astronomy. All other things being equal, sources which are farther away contribute less gravitational wave amplitude. The chirp mass ($\mathcal{M}_c$) is a combination of the masses of the two binary objects where $\mathcal{M}_c = (m_1 m_2)^{3/5}(m_1 + m_2)^{-1/5}$, and serves a similar role in gravitational wave analysis to that of reduced mass in analysis of atomic systems.

The source of the $\dot{f}$ parameter comes from two categories of phenomena. The UCB is emitting gravitational waves, which will tend to increase its orbital frequency ($\dot{f}_{gr}$). There

also may be some mass transfer, or other feedback occurring between the two orbiting objects. Depending on the exact process involved, this may increase or decrease the orbital frequency.

As an example, AM CVn systems are thought to have mass transfer that can decrease or increase the orbital frequency over time [17]. This astrophysical contribution from all sources of frequency change other than ideal gravitational wave emission is $\dot{f}_{astro}$. The parameters used in this work treat $\dot{f} = \dot{f}_{gr} + \dot{f}_{astro}$ to keep generality with these interacting binary sources that we expect to appear in LISA observations.

Prior to this project `GBMCMC` used $\{f_0, \theta, \phi, \iota, \psi, \phi_0, \mathcal{A}, \dot{f}\}$ as the search space for each gravitational wave source. The parameter space in this work must necessarily include D_L and $\mathcal{M}_c$, as well as $\dot{f}_{astro}$ in order to incorporate physically motivated priors in the search. The transform between $\{..., \mathcal{A}, \dot{f}\}$ and $\{..., D_L, \mathcal{M}_c, \dot{f}_{astro}\}$ parameterizations is possible via the following relationships[15, 9]:

$$\mathcal{A}(f_0, D_L, \mathcal{M}_c) = \frac{2}{D_L}\left(\frac{G\mathcal{M}_c}{c^2}\right)^{5/3}\left(\frac{\pi f_0}{c}\right)^{2/3} \tag{2.1}$$

$$\dot{f}(f_0, \mathcal{M}_c, \dot{f}_{astro}) = \frac{96\pi^{8/3}}{5}\left(\frac{G\mathcal{M}_c}{c^3}\right)^{5/3}(f_0)^{11/3} + \dot{f}_{astro} \tag{2.2}$$

This physical parameter space is degenerate on account of the additional degree of freedom in the $\{..., D_L, \mathcal{M}_c, \dot{f}_{astro}\}$ representation. The left side of figure 2.1 shows this degeneracy as curves of constant amplitude plotted in $\mathcal{M}_c$ and D_L, at a frequency of $f_0 = 3.4$mHz. While the right-hand graph shows a similar relationship between chirp mass and $\dot{f}_{GR}$ for a range of principle frequencies in the LISA band.

The left side of figure 2.1 is the essential distance/luminosity relationship for UCB sources. The right-hand graph illustrates the speed at which the orbital period of an ideal UCB changes for a given chirp mass and initial frequency independent of amplitude. It is expected from general relativity that detached binaries will exactly meet this criteria, while binaries with mass transfer or feedback dynamics will differ from this relationship. The unit choice on the right-hand graph places some limits on how precisely the frequency of a UCB must be

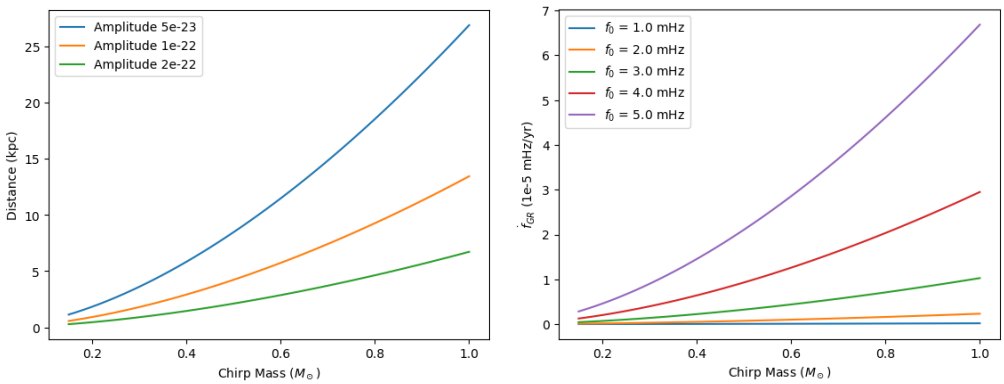

Figure 2.1: Diagram of distance/mass relationship for constant amplitude (left) and $\dot{f}_{GR}$ vs mass for constant frequency (right).

known in order to begin to apply such an analysis. In short, the frequency must be known well enough and over a long enough observation time that its derivative can be characterized, and this is less constraining for higher frequency sources, and higher chirp masses.

In terms of implementation, this modification to allow `GBMCMC` to operate in a different parameter space was time-consuming because it was necessary to maintain compatibility with the use of a second frequency derivative parameter ($\ddot{f}$) as an optional mode in `GBMCMC`, as well as maintain compatibility between `GBMCMC` and `GLASS`. The implementation therefore had to cover 4 possible parameter sets (two described above, plus each with $\ddot{f}$). This in turn required `GBMCMC` and `GLASS` code to use named constants as parameter referents rather than integers. Because the code is written in C, this change required additional bookkeeping to maintain performance[1]. It also included the error-prone process of modifying every parameter access across the entire codebase[2].

[1] `gb_params.h` contains this bookkeeping system.

[2] All of the modifications to `GBMCMC` and `GLASS` used in this work can be found at `https://github.com/mtauraso/ldasoft/tree/Tauraso-Thesis`

Chapter 3

TRANS-DIMENSIONAL MCMC FOR UCBS

This is a review of the broad strokes of the detailed balance for a Markov Chain Monte Carlo process, focusing on the Metropolis-Hastings algorithm, and its trans-dimensional generalization, the Reversible Jump algorithm[5]. The focus is on drawing a connection between the gravitational wave source parameter finding problem as implemented in `GBMCMC` and the detailed balance condition that is at the heart of why MCMC algorithms work. The majority of this explanation can be found in a more measure-theoretic form in Chapter 6 of [6] as well as in Section 3 of [5]. I have attempted to translate the explanation to the language of multi-variable calculus so that the workings of it are accessible without understanding measure theory. I have also attempted to make explicit the link between the Bayesian formulation typical in the sciences, and the abstract notation typical in the mathematical literature.

3.1 Detailed Balance

What we're trying to find is a set of parameters such that the gravitational wave signal defined by those parameters best matches the gravitational wave signal from the LDC data. In order to do this, we want to get a probability distribution defined over the parameter space, which peaks at the most likely parameters. Our parameters are a vector $\mathbf{x} =< f_0, \theta, \phi, ... >$. This vector is assumed to live in a flat Cartesian phase space such that $d\mathbf{x} \equiv df_0 d\theta d\phi ...$.

Given the observed signal $\mathbf{d}$, the probability distribution that peaks at our ideal parameter values is the posterior $\pi(\mathbf{x}) \equiv P(\mathbf{x}|\mathbf{d})$. We don't care about the normalization of this function. Any constant times $\pi(\mathbf{x})$ will do for our purposes, because we most care about where in the phase space of $\mathbf{x}$ it has a maximum and how wide that maximum is.

Bayes' theorem tells us that we can essentially reverse the order of a conditional probability expression at the cost of a fraction between the two probabilities on either side of the bar. For our posterior,

$$\pi(\mathbf{x}) = P(\mathbf{x}|\mathbf{d}) = P(\mathbf{d}|\mathbf{x})\frac{P(\mathbf{x})}{P(\mathbf{d})}. \tag{3.1}$$

When our data d are known, $P(\mathbf{d})$ is a constant factor, so by using our constant factor flexibility in π we can simply redefine $\pi(\mathbf{x}) \equiv P(\mathbf{d}|\mathbf{x})P(\mathbf{x})$. We can equivalently view this function as the product of the prior for our parameters, $P(\mathbf{x})$ and the likelihood $L(\mathbf{x}) \equiv P(\mathbf{d}|\mathbf{x})$. The calculation of the likelihood for a given gravitational wave signal given some parameters is outlined in [14].

We're going to approximate $\pi(\mathbf{x})$ using a Markov process. At each step of the process we generate a new set of parameters $\mathbf{x}'$ using the prior set of parameters $\mathbf{x}$ as input. This process will ultimately generate a chain of values. The chains created by this process have a recurring probability distribution $Q(\mathbf{x}, \mathbf{x}')$, defined as the probability of $\mathbf{x}'$ appearing given $\mathbf{x}$ in the previous step of the chain. $Q(\mathbf{x}, \mathbf{x}')$ is called the kernel of the Markov process.

We want $\pi(\mathbf{x})$ to be a fixed point of this process, such that if you start with a set of chain samples having distribution $\pi(\mathbf{x})$, continuing the process will keep the samples having the same distribution. We can think of this as choosing an $\mathbf{x}$ value from $\pi(\mathbf{x})$, and insisting that we have the same probability to get some $\mathbf{x}'$ as we would to do the reverse, drawing $\mathbf{x}'$ from $\pi(\mathbf{x}')$ to get get $\mathbf{x}$ as the next value in the chain. This detailed balance of the probabilities to get additional data points is a key property that allows Markov chains to eventually converge on the posterior distribution.

We can strengthen this notion of detailed balance beyond individual points, to any arbitrary region of parameter space. Given some regions A, B in $\mathbb{R}^n$ such that $\mathbf{x} \in A$ and $\mathbf{x}' \in B$ we can write a detailed balance condition as follows:

$$\int_A \int_B \pi(\mathbf{x})Q(\mathbf{x}, \mathbf{x}')\, d\mathbf{x}' d\mathbf{x} = \int_B \int_A \pi(\mathbf{x}')Q(\mathbf{x}', \mathbf{x})\, d\mathbf{x} d\mathbf{x}' \tag{3.2}$$

In the Metropolis-Hastings method of MCMC, the $\mathbf{x}'$ is generated by drawing values from a proposal distribution, with probability $q(\mathbf{x}, \mathbf{x}')$ and then accepting that proposal with a probability $\alpha(\mathbf{x}, \mathbf{x}')$. Considering both the cases where $\mathbf{x}'$ is accepted and rejected, the kernel of a Metropolis-Hastings MCMC can be written as

$$Q(\mathbf{x}, \mathbf{x}')d\mathbf{x}' = q(\mathbf{x}, \mathbf{x}')\alpha(\mathbf{x}, \mathbf{x}')d\mathbf{x}' + I(\mathbf{x} \in d\mathbf{x}')s(x) \tag{3.3}$$

$$s(x) = \int_C (1 - \alpha(\mathbf{x}, \mathbf{x}'))q(\mathbf{x}, \mathbf{x}')d\mathbf{x}' \tag{3.4}$$

The first term in equation 3.3 is the probability that $\mathbf{x}'$ is chosen from q and then accepted. The second term together with equation 3.4 describe the probability that the new sample is rejected. In this notation we are considering $d\mathbf{x}'$ as a tiny region in $\mathbb{R}^n$. The indicator function, $I(x \in d\mathbf{x}')$ takes the value 1 whenever x is in $d\mathbf{x}'$, and zero otherwise. In equation 3.4, C is the set of all possible parameters $\mathbf{x}'$ that could have been drawn and were discarded.

Using equation 3.3 we can now rewrite the detailed balance condition in terms of the functions that define the Metropolis-Hastings process as

$$\int_A \pi(\mathbf{x}) \int_B q(\mathbf{x}, \mathbf{x}')\alpha(\mathbf{x}, \mathbf{x}')\, d\mathbf{x}d\mathbf{x}' + \int_{A\cap B} \pi(\mathbf{x})s(\mathbf{x})d\mathbf{x} = \int_B \pi(\mathbf{x}') \int_A q(\mathbf{x}', \mathbf{x})\alpha(\mathbf{x}', \mathbf{x})\, d\mathbf{x}d\mathbf{x}' + \int_{B\cap A} \pi(\mathbf{x}')s(\mathbf{x}')d\mathbf{x}'. \tag{3.5}$$

The integrals over $A \cap B$ on both sides of equation 3.5 cancel, leaving us with a more compact form describing how the Metropolis-Hastings functions must be related for the algorithm to balance.

$$\int_A \int_B \pi(\mathbf{x})q(\mathbf{x}, \mathbf{x}')\alpha(\mathbf{x}, \mathbf{x}')\, d\mathbf{x}d\mathbf{x}' = \int_A \int_B \pi(\mathbf{x}')q(\mathbf{x}', \mathbf{x})\alpha(\mathbf{x}', \mathbf{x})\, d\mathbf{x}d\mathbf{x}' \tag{3.6}$$

The α which satisfies this balance for an arbitrary proposal distribution q and posterior π can be found by equating the integrands to give the following:

$$\frac{\alpha(\mathbf{x},\mathbf{x}')}{\alpha(\mathbf{x}',\mathbf{x})} = \frac{\pi(\mathbf{x}')q(\mathbf{x}',\mathbf{x})}{\pi(\mathbf{x})q(\mathbf{x},\mathbf{x}')} \tag{3.7}$$

In most implementations it makes sense to define the acceptance ratio α as a probability with a maximum value of 1. In the case where both sides of equation 3.7 are less than 1, $\alpha(\mathbf{x},\mathbf{x}')$ is simply the right hand side, and $\alpha(\mathbf{x}',\mathbf{x}) = 1$. When both sides are greater than 1, then $\alpha(\mathbf{x},\mathbf{x}') = 1$ and $\alpha(\mathbf{x}',\mathbf{x})$ is the reciprocal of the right hand side. This can be written compactly in the canonical fashion:

$$\alpha(\mathbf{x},\mathbf{x}') = min\left\{1, \frac{\pi(\mathbf{x}')q(\mathbf{x}',\mathbf{x})}{\pi(\mathbf{x})q(\mathbf{x},\mathbf{x}')}\right\}. \tag{3.8}$$

Expanding $\pi(\mathbf{x})$ into our prior and likelihood functions from earlier gives the form:

$$\alpha(\mathbf{x},\mathbf{x}') = min\left\{1, \frac{P(\mathbf{x}')L(\mathbf{x}')q(\mathbf{x}',\mathbf{x})}{P(\mathbf{x})L(\mathbf{x})q(\mathbf{x},\mathbf{x}')}\right\}. \tag{3.9}$$

3.2 Transdimensional Jumps

We have so far assumed that $\mathbf{x}$ and $\mathbf{x}'$ are concretely the parameters for a single galactic binary, and using equation 3.9 it is possible to implement a Metropolis-Hastings MCMC algorithm to find the parameters of a single galactic binary from its gravitational waves. However, LISA data has multiple overlapping gravitational wave signals. We therefore need a way to extend this technique to multiple sources.

We are going to keep the same Markov chain idea; however, we are going to permit jumps that add an entirely new vector of parameters $\mathbf{y} =< f_0, \theta, \phi, ... >$ for a second source. We are also going to permit jumps which remove a vector of parameters. In similar fashion to the single-source MCMC, we'd like for these sorts of jumps to also satisfy the detailed balance equation.

The mathematical argument that lead to equation 3.8 from the detailed balance condition, works for any number of dimensions, and even works if A and B do not have the same

dimension, so long as they are subsets of $\mathbb{R}^n$. We can exploit this to modify the Metropolis-Hastings idea to create jumps between dimensions.

The rule for jumping dimensions is that we generate a random vector $\mathbf{u}$ from a probability density $g(\mathbf{u})$. The new state is then generated by a function $h(\mathbf{x}, \mathbf{u}) = \mathbf{x}'$. We then accept this $\mathbf{x}'$ based on an acceptance ratio $\alpha(\mathbf{x}, \mathbf{x}')$, similar to the old Metropolis-Hastings process. The process of generating a new state vector must be reversible by generating a random vector[1] $\mathbf{u}'$ from a probability density $g'(\mathbf{u}')$, and using a function $h'(\mathbf{x}', \mathbf{u}') = \mathbf{x}$ to reconstruct the original state.

We also require that the transformation functions h and h' be differentiable, and that the total dimension of $(\mathbf{x}, \mathbf{u})$ and $(\mathbf{x}', \mathbf{u}')$ be equal. Given those constraints, we can write the detailed balance equation again. Equation 3.10 is the same as 3.2, but the brackets associate regions A and B with $\mathbf{x}$ and $\mathbf{x}'$ respectively to help keep track of the dimension of each space.

$$\int_A \pi(\mathbf{x}) \left[\int_B Q(\mathbf{x}, \mathbf{x}') \, d\mathbf{x}' \right] d\mathbf{x} = \int_B \pi(\mathbf{x}') \left[\int_A Q(\mathbf{x}', \mathbf{x}) \, d\mathbf{x} \right] d\mathbf{x}' \tag{3.10}$$

Following the same logic as before, both sides of equation 3.10 contain a term describing rejection of the sample, which cancel one another. Considering again equation 3.5, the cancelled terms are both integrals over $A \cap B$. Even now that A and B have different dimension, the integral over their intersection still cancels. Following the earlier logic through equation 3.6 gives us the detailed balance in terms of the proposal probability distribution in a Metropolis-Hastings framework.

$$\int_A \pi(\mathbf{x}) \left[\int_B q(\mathbf{x}, \mathbf{x}') \alpha(\mathbf{x}, \mathbf{x}') \, d\mathbf{x}' \right] d\mathbf{x} = \int_B \pi(\mathbf{x}') \left[\int_A q(\mathbf{x}', \mathbf{x}) \alpha(\mathbf{x}', \mathbf{x}) \, d\mathbf{x} \right] d\mathbf{x}' \tag{3.11}$$

Note that the bracketed integrals in equation 3.11 are defined in terms of $\mathbf{x}$ and $\mathbf{x}'$; however, the h and g functions allow us to re-write them in terms of $\mathbf{u}$ and $\mathbf{u}'$. In equation 3.12 it is assumed that $\mathbf{u} \in U$, $\mathbf{u}' \in V$, and $U, V \subset \mathbb{R}^n$.

---

[1] In GLASS, and in the motivating example of adding a source $\mathbf{y}$, it is common for $\mathbf{u}'$ to be of zero length; however, it is kept here for generality.

$$\int_A \pi(\mathbf{x}) \left[\int_U g(\mathbf{u}) \alpha(\mathbf{x}, \mathbf{x}') d\mathbf{u} \right] d\mathbf{x} = \int_B \pi(\mathbf{x}') \left[\int_V g'(\mathbf{u}') \alpha(\mathbf{x}', \mathbf{x}) d\mathbf{u}' \right] d\mathbf{x}'$$

$$\int_A \int_U \pi(\mathbf{x}) g(\mathbf{u}) \alpha(\mathbf{x}, \mathbf{x}') \, d\mathbf{x} \, d\mathbf{u} = \int_B \int_V \pi(\mathbf{x}') g'(\mathbf{u}') \alpha(\mathbf{x}', \mathbf{x}) \, d\mathbf{x}' d\mathbf{u}' \tag{3.12}$$

The dimension match required earlier on $(\mathbf{x}, \mathbf{u})$ and $(\mathbf{x}', \mathbf{u}')$ ensures that the two spaces we're integrating over in equation 3.12 have the same dimension. We can therefore equate the integrands by a change of variables. This change of varibles will use the determinant of the Jacobian, which is ultimately defined by the derivatives of the h and h' functions. Using this change of variables, we can write a similar ratio as before to isolate the acceptance ratio α.

$$\frac{\alpha(\mathbf{x}, \mathbf{x}')}{\alpha(\mathbf{x}', \mathbf{x})} = \frac{\pi(\mathbf{x}') g'(\mathbf{u}')}{\pi(\mathbf{x}) g(\mathbf{u})} \left| \frac{\partial(\mathbf{x}', \mathbf{u}')}{\partial(\mathbf{x}, \mathbf{u})} \right| \tag{3.13}$$

Similar logic as before gives the acceptance ratio for a reversible jump MCMC in terms of our prior $P(\mathbf{x})$ and likelihood $L(\mathbf{x})$ as:

$$\alpha(\mathbf{x}, \mathbf{x}') = min \left\{ 1, \frac{P(\mathbf{x}') L(\mathbf{x}') g'(\mathbf{u}')}{P(\mathbf{x}) L(\mathbf{x}) g(\mathbf{u})} \left| \frac{\partial(\mathbf{x}', \mathbf{u}')}{\partial(\mathbf{x}, \mathbf{u})} \right| \right\} \tag{3.14}$$

In `GBMCMC` this reversible jump acceptance is used for the birth/death and split/merge moves, which each add or subtract a source from the model under consideration. `GBMCMC` makes several simplifying assumptions within this general formulation. $\mathbf{u}'$ has only a single uniform variable u', which is only used to select a source to remove. This choice means $\mathbf{u}$ must be one larger than $\mathbf{y}$ to balance dimensions. For convenience I will keep $\mathbf{u}$ and $\mathbf{y}$ the same size and simply call this additional number u' in both contexts[1].

In the case of adding a source, $\mathbf{x}'$ is simply the original parameters $\mathbf{x}$ and the newly added parameters $\mathbf{y}$. The new parameters $\mathbf{y}$ are generated exclusively from the random numbers $\mathbf{u}$.

[1] `GBMCMC` doesn't bother generating u' in the creation case, because the number would be thrown away, and never contribute to the calculation. It is only important to track here because it's derivative appears in the Jacobian

Therefore the functions h and h' serve only to scale $\mathbf{u}$ to astrophysical parameter ranges. As a result of this simplicity, the only meaningful derivatives in the Jacobian are $\frac{\partial \mathbf{y}}{\partial \mathbf{u}}$, which are simply a diagonal matrix of scale factors. The Jacobian determinant simplifies as follows:

$$\left|\frac{\partial(\mathbf{x},\mathbf{y},u')}{\partial(\mathbf{x},\mathbf{u},u')}\right| = \begin{vmatrix} \frac{\partial \mathbf{x}}{\partial \mathbf{x}} & \frac{\partial \mathbf{x}}{\partial \mathbf{u}} & \frac{\partial \mathbf{x}}{\partial u'} \\ \frac{\partial \mathbf{y}}{\partial \mathbf{x}} & \frac{\partial \mathbf{y}}{\partial \mathbf{u}} & \frac{\partial \mathbf{y}}{\partial u'} \\ \frac{\partial u'}{\partial \mathbf{x}} & \frac{\partial u'}{\partial \mathbf{u}} & \frac{\partial u'}{\partial u'} \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & \frac{\partial \mathbf{y}}{\partial \mathbf{u}} & 0 \\ 0 & 0 & 1 \end{vmatrix} = \left|\frac{\partial \mathbf{y}}{\partial \mathbf{u}}\right| \tag{3.15}$$

If we define a proposal probability density $q_{new}(\mathbf{y})$ for the new source parameters, and insist it be normalized, then we can equate it with the normalization condition for g giving us $\int g(\mathbf{u})d\mathbf{u} = 1 = \int q(\mathbf{y})d\mathbf{y}$. Changing variables and equating these integrands allows us to see that

$$q(\mathbf{y})\left|\frac{\partial \mathbf{y}}{\partial \mathbf{u}}\right| = g(\mathbf{u}) \tag{3.16}$$

This result can be substituted into the RJMCMC acceptance ratio from earlier. If we also rewrite the prior $P(\mathbf{x}') = P(\mathbf{x})P(\mathbf{y})$ we can recover the form used by `GBMCMC` (formula 11 in [14]).

$$\alpha(\mathbf{x},\mathbf{x}') = min\left\{1, \frac{L(\mathbf{x}')P(\mathbf{y})}{L(\mathbf{x})q(\mathbf{y})}\right\} \tag{3.17}$$

Chapter 4

PROPOSALS

`GBMCMC` relies on both fixed and trans-dimensional MCMC steps to choose a source, relying on the same proposals to both propose new sources for consideration in the trans-dimensional (RJMCMC) steps, and to propose changes to existing sources in fixed-dimensional MCMC steps. For proposal methods which are used in both a trans-dimensional and a fixed context, the draw procedure is approximately the same; however, the assembly of the acceptance ratio differs slightly between the two cases. Table 4.1 contains an overview of the proposal cocktail in `GBMCMC`. Each proposal has a probability to be chosen during a particular MCMC or RJMCMC iteration. It should be noted as well that 100 MCMC iterations occur for every RJMCMC iteration in `GBMCMC`.

Proposal	MCMC probability	RJMCMC probability	Modification scope
uniform draw	0.1	0.2	large
F-statistic draw [3, 7]	0.2	The rest	small
GMM/COV Draw (optional)	0.2	0.2	trivial
Fisher Matrix [11]	The rest	0.0	large
FM shift	0.1	0.0	trivial
ψ-ϕ_0 jump	0.2	0.0	trivial

Table 4.1: Proposals by weight and scope of modification to implement distance prior

Lines in table 4.1 where the weight is shown as "The rest" indicate that at runtime that proposal has a probability to be run of 1 minus the sum of all proposals in the column. The set

of proposals cannot be known until runtime because the Gaussian mixture model/covariance matrix (GMM/COV) draw proposal is not always used. The GMM/COV draw proposal uses a earlier run of `GBMCMC` to inform a subsequent run and enhance convergence[14, 10]. It is important to note that the F-statistic draw proposal and the Fisher matrix proposal have much of the weight and are the most critical proposals to the search.

The overall change in parameter space required several proposals to be updated. The GMM/COV draw, FM shift, and ψ-ϕ_0 jump proposals did not require significant modification because they function similarly in both parameter spaces; however the uniform and F-statistic draws needed to be modified to draw from the 3-D volume defined by the new parameter space. The Fisher Matrix became singular in the new parameter space, so required major modification.

4.1 Uniform and F-Statistic Draw

The F-statistic and uniform draw proposals together cover 80-100% of RJMCMC proposals in any run of `GBMCMC`. They both rely on uniformly drawing all parameters; however the F-statistic draw takes the additional step of rejecting some samples based on an F-statistic evaluated over the data as described in [3]. The uniform draw used by both of these proposals is correct for chirp mass ($\mathcal{M}_c$), and astrophysical frequency change ($\dot{f}_{astro}$); however, applied to distance (D_L) it creates some problems.

Drawing distance uniformly would preferentially propose locations that are not uniformly distributed over the search volume in physical space. `GBMCMC` would preferentially search the outer boundaries of the bounding sphere where it is sensitive, and would have a very small chance to find sources in the galaxy.

The solution to this issue is a distance draw probability density made up of two probability density functions defined over the spherical search space. The two density functions are $f(\mathbf{r})$, which has constant density per unit volume, and $g(\mathbf{r})$ which has constant density over coordinate r. $\mathbf{r} = \langle D_L, \theta, \phi \rangle$ is a point within the search space, a sphere of radius R centered on earth. The combined probability density function $h(\mathbf{r})$ is defined in equation 4.1, with

the parameter $a \in [0, 1)$[1] defining the relative weight of the two functions.

$$f(\mathbf{r}) = \frac{1}{4\pi}\frac{3}{R^3} \qquad g(\mathbf{r}) = \frac{1}{4\pi}\frac{1}{r^2 R} \qquad h(\mathbf{r}) = (1-a)f(\mathbf{r}) + (a)g(\mathbf{r}) \tag{4.1}$$

For the purposes of drawing only r coordinate, the common factor of $(4\pi)^{-1}$ can be ignored because it is evaluated by the draw functions for ϕ and θ, which are themselves drawn uniformly over solid angle. Therefore the code that draws D_L must only evaluate the radial factor of $h(\mathbf{r})$. This factor can be written entirely in terms of the coordinate r as:

$$h_{radial}(r) = \frac{1}{R}\left[(1-a)\frac{3}{R^2} + \frac{a}{r^2}\right] \tag{4.2}$$

$h_{radial}(r)$ has normalization condition $\int_0^R h_{radial}(r) r^2 dr = 1$, which follows from $f(\mathbf{r})$ and $g(\mathbf{r})$ being normalized inside a sphere of radius R about the origin. Considering $h_{radial}(r)r^2$ as a 1-dimensional probability density function (PDF) in coordinate r means that drawing a location in space from $h(\mathbf{r})$ can be reduced to isotropically choosing a sky location, and then drawing r according to[2]

$$h_{draw}(r) = \frac{1}{R}\left[(1-a)\frac{3r^2}{R^2} + a\right] \tag{4.3}$$

Equation 4.3 demonstrates the need for two functions in the initial 3-dimensional proposal density. Recall that the first term in h_{draw} arises from the uniform-in-volume density function $f(\mathbf{r}) = \frac{1}{4\pi}\frac{3}{R^3}$. If we set $a = 0$, only $f(\mathbf{r})$ will contribute probability, and $h_{draw}(r) = 3r^2/R^3$. For r coordinate very near the earth, this proposal distribution lacks support. The result for $a = 0$ is a cutoff in the search where near sources are not identified as being near earth in the posterior. Introducing $g(\mathbf{r})$ is a remedy that provides support in the proposal distribution that allows the MCMC to explore sources near earth.

---

[1] a in this section is defined by `DIST_UNI_DRAW` in `GBMCMC` and is set to $a = 0.2$. Section 6.3.1 contains discussion of this choice.

[2] $ln(h_{draw}(r))$ is calculated by `distance_draw_logP()` in `GalacticBinaryProposal.c`

In practice, h_{draw} is used in a rejection sampling algorithm where each iteration compares the PDF value to a random value in the range $[0, 1)$, accepting when the PDF is less than the random value. The rejection rate of this algorithm is minimized when the PDF's maximum value is 1. The maximum value h_{draw} reaches is $\frac{3-2a}{R}$. Therefore, the rejection sampling algorithm uses equation 4.4[3] to minimize the rejection rate.

$$P_{accept}(r) = \frac{1}{3-2a}\left[(1-a)\frac{3r^2}{R^2} + a\right] \tag{4.4}$$

4.2 Fisher Matrix Proposal

The Fisher matrix proposal in `GBMCMC` works by proposing jumps in parameter space with a randomly selected distance along the eigenvectors of the Fisher information matrix of the likelihood function. Near the correct parameter values, it can be shown that the likelihood surface ought to take on a multivariate Gaussian character, which is parameterized by the Fisher information matrix[4]. The exact implementation is approximate, and relies on correctly scaled jumps along the eigenvectors of the Fisher information matrix. The result is an efficient MCMC proposal near likelihood maxima[11].

In the new parameter search space $\{..., D_L, \mathcal{M}_c, \dot{f}_{astro}\}$, the nine search parameters are now a degenerate basis of the 8-dimensional space that defines the gravitational wave template for an individual source. The Fisher matrix is therefore singular, and it's eigenvectors are no longer a good approximation of this multivariate Gaussian in all possible directions. The remedy is to calculate the Fisher matrix three times, once for each possible pair of the three newly introduced parameters. When the distance prior is in use, the Fisher matrix proposal computes jumps in the same manner, but has a uniform probability to choose among the three Fisher matrices.

A motivation for this scheme can be seen by considering Fisher information as the natural metric of an information manifold [18]. Any possible parameterization of a gravitational wave

[3] $P_{accept}(r)$ is calculated in `distance_draw_P() in GalacticBinaryProposal.c`

is simply a different chart over this same manifold. So long as our choice of chart does not have a singularity near to the point the MCMC is considering, one chart is as good as the next. In practice the code detects singular matrices and does not use this proposal mechanism in areas of parameter space where one of these charts is singular.

Chapter 5
DISTANCE PRIOR

This section gives an overview of the construction of the distance prior. The core idea is to extend `GBMCMC`'s notion of a prior on the sky to a prior on both sky location and distance for these sources.

5.1 Sky Prior

`GBMCMC` originally incorporates a prior on the $\{\theta, \phi\}$ sky location. This prior is based on an exponential disk model of the galaxy, which is computed and projected on to the sky using an MCMC integration scheme. A uniform contribution is then added to every sky direction so there is prior support for every direction in the sky[14]. The distance prior reuses this galaxy model, and the overall approach, with some small modifications described in the next section. The probability density function for the galaxy model in galactic (x, y, z) coordinates:

$$P_{galaxy}(x,y,z)^1 = P_{bulge}e^{-r^2/R_{bulge}^2} + (1 - P_{bulge})\frac{e^{-u/R_{disk}}}{\cosh^2(z/Z_{disk})} \tag{5.1}$$

In equation 5.1, $r = \sqrt{x^2 + y^2 + z^2}$ is the distance to the origin, and $u = \sqrt{x^2 + y^2}$. In calculating the sky prior, a lattice of 200x200 bins is constructed over the full range of $cos(\theta) \in [-1, 1]$ and $\phi \in [0, 2\pi)$ such that each bin contains the same solid angle. An MCMC integration scheme is used to evaluate P_{galaxy} at each lattice point.

The MCMC integration searches a volume of 100x100x32 kpc centered on the center of the Milky Way, using $P_{galaxy}(x, y, z)$ as the target function for the MCMC. The resulting

[1]The constants $P_{bulge} = 0.25$, $R_{bulge} = 0.8$kpc, $R_{disk} = 2.5$kpc, and $Z_{disk} = 0.4$kpc are tunable parameters of the galaxy density model, and were not altered for this work.

chain of (x, y, z) values is then mapped to $(\phi, cos(\theta))$ bins, corresponding to where each value appears in the sky from earth. The counts in the $(\phi, cos(\theta))$ bins are then normalized, and the uniform contribution is added to produce a lookup table. In the total sky prior the uniform contribution has weight 0.1, and the galaxy model portion has weight 0.9.

5.2 Distance Prior

The distance prior is implemented using the same P_{galaxy} function used for the sky prior, combined with an isotropic probability density to give prior support over all sky angles. Unlike the MCMC scheme in use for the sky prior, the distance prior is calculated directly for every spherical volume element[2] in a sphere around earth with 200 subdivisions in each of D_L, $cos(\theta)$, and ϕ. The maximum radius of this bounding sphere is calculated from the LISA noise sensitivity curve in a manner detailed in section 5.3. Integration inside the same bounding sphere is used to normalize the overall probability density.

The distance prior also has an isotropic component analogous to the uniform component present in the sky prior. The motivation for this is similar, to provide some prior support to find sources not in the galactic plane. The isotropic component of the prior uses the volume probability density function:

$$P_{uni} = (r^2 + a^2)^{-1} \tag{5.2}$$

where $a = 10kpc$ was chosen to reduce the probability density near $r = 0$ while providing tails which are uniformly distributed over sky solid angle and coordinate r. Uniform density over coordinate r was chosen to prevent prior support near the outer boundary of the bounding sphere from overwhelming support in the combined prior for sources in the galactic disk.

The distance prior is a weighted sum of the galaxy based prior and the uniform prior,

[2] In `GBMCMC` code the distance prior is often called the volume prior because it is defined over every spherical volume element in the bounding sphere. The distance prior is enabled with the `--volume-prior` command line flag.

with 0.1 weight given to the uniform prior and 0.9 given to the galaxy based prior[3]. The prior density is held in memory as a lookup table[4] while `GBMCMC` runs, similar to the sky prior.

Early versions of the distance prior used the sky prior's MCMC integration scheme. This was too sparse over a 3-dimensional volume, with many volume elements receiving zero visits from the MCMC. Increasing the iterations of the MCMC was more computationally intensive than simply calculating the prior in each volume element independently.

Cartesian volume elements and spherical volume elements centered on the galactic center were also tried. Both of these approaches suffered from anisotropy introduced by the volume elements themselves, which then spoiled sky location posteriors.

5.3 Search Volume

The radius of the search volume is based on analysis of the sensitivity of the LISA instrument at different frequencies, and the principle that the MCMC search should not consider sources that are so far away from earth, or with such low mass that they cannot be identified by the LISA instrument due to noise. Making such an assessment involves combining characterizations of the LISA instrumental noise, confusion noise arising from UCBs below the detection threshold, as well as prior assumptions about the maximum chirp mass of UCBs. `GBMCMC` uses a noise model as part of the overall MCMC pipeline, and the analytic noise model configured for LDC Radler data is used for the explanation below.

Since `GBMCMC` runs on a narrow frequency band, first it estimates the minimum noise level within that band using the existing analytic noise model. Then it calculates a gravitational wave amplitude $\mathcal{A}$ corresponding to an SNR of 5 given the observational time under analysis. Finally it calculates how far away a UCB source could be and still generate a signal with amplitude $\mathcal{A}$. Equation 5.3 is used with $\mathcal{M}_c = 1M_\odot$, the maximum of the prior on chirp mass,

[3]The weighting of this component of the prior is set by `GalacticBinaryPrior.h:VOL_PRIOR_UNI`.

[4]Running `GBMCMC` with the `--verbose` option will dump the raw sky or volume prior lookup table into the current working directory.

and at the maximum frequency in the analysis band to determine the maximum distance[5].

$$D_L = \frac{2}{\mathcal{A}} \left(\frac{G\mathcal{M}_c}{c^2} \right)^{5/3} \left(\frac{\pi f_{max}}{c} \right)^{2/3} \tag{5.3}$$

The relationship between this maximum distance along with frequency and observation time is shown in figure 5.1. The "knee" in the power law relationship shown is caused by the trough in the LISA noise curve around 5 mHz. Greater noise above 5mHz limits the distance that LISA can hear this prototypical loudest source. The waviness in the power law portion of figure 5.1 below 5 mHz is primarily the result of the confusion noise from non-resolvable UCB sources. The tick upward at the high end of frequency is a resonance of the interferometer model in use.

The curves in figure 5.1 are somewhat involved to calculate because they incorporate one of many possible models of LISA noise. Below 6 mHz the following approximate power law expression can be used to derive the distance with fewer steps.

$$R_{max} = 70\,(T_{obs} + 1) \left(\frac{f}{6\,\mathrm{mHz}} \right)^{1.4} \left(\frac{\mathrm{kpc}}{\mathrm{yr}} \right) \tag{5.4}$$

The original motivation for finding this approximation was to estimate the size of the bounding sphere for the volume prior in a computationally efficient manner; however, it is too inaccurate for this purpose. Underestimating the bounding sphere radius risks not exploring spatial volume that could contain observable sources. Overestimation wastes computational resources exploring volume where `GBMCMC` mathematically cannot find any sources due to its settings and internal noise model. This wasted time was expected to compound at larger distances. This approximation also lacks robustness to the various ways that `GBMCMC` can be configured with a novel model of LISA noise. I therefore present equation 5.4 only as a path to an order of magnitude estimate for the cutoff distance.

---

[5]This calculation occurs in `GalacticBinaryPrior.c:set_search_volume()`

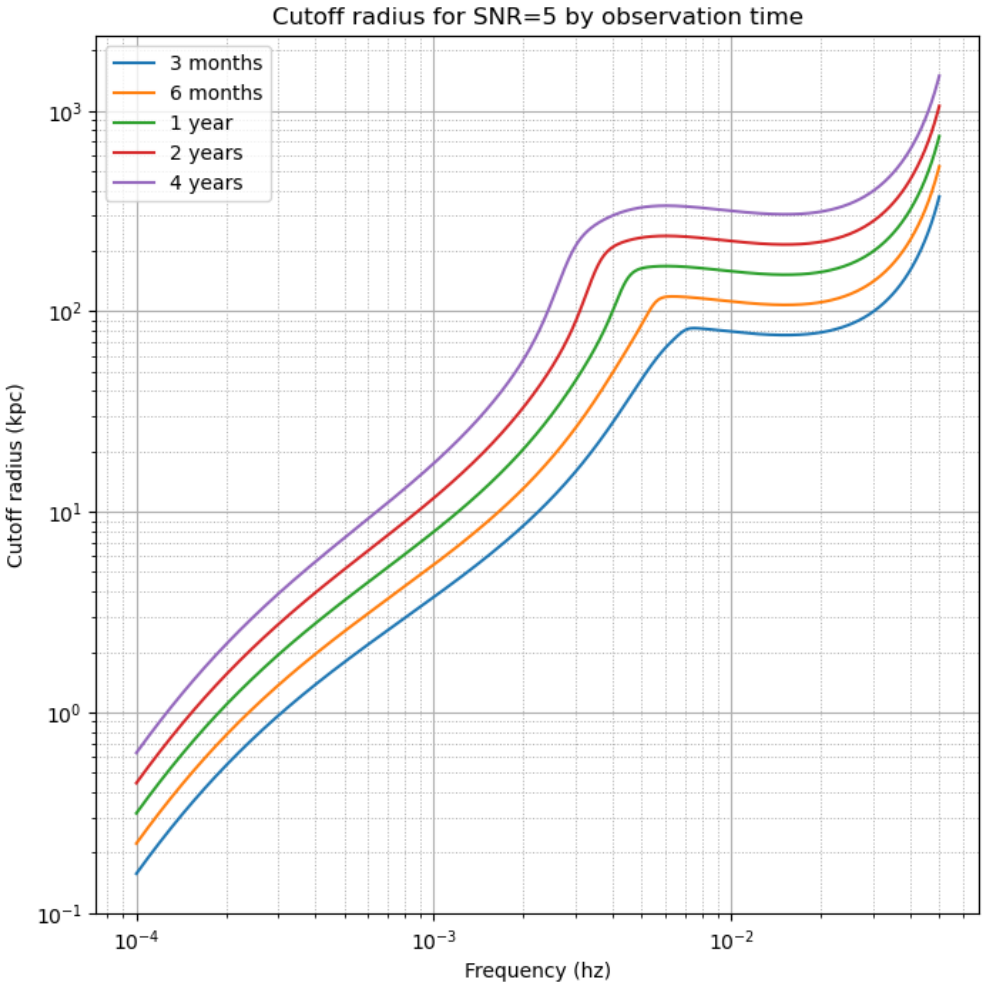

Figure 5.1: Cutoff radius versus frequency for the bounding sphere of the volume prior showing dependence on frequency and observation time.

Chapter 6

RESULTS

In understanding these results it is important to note that `GBMCMC` uses a parallel tempering scheme, where 'hot' and 'cold' chains are calculated concurrently. The chains used for the results below are the coldest chains, which are expected to be the best estimators of UCB parameters. `GBMCMC` uses the number of central processing units (CPUs) available at runtime to determine how many parallel chains will be run. For all of the results below 5 AMD EPYC 9654 CPUs were used for each run on the UW Hyak compute cluster.

6.1 Detailed Balance Testing

As a consistency check of the new distance proposal it is necessary to understand empirically whether or not it is biasing the search. `GBMCMC` can be run in a mode[1] where likelihood is set equal to 1. In this mode the posteriors of the MCMC search are only a result of the priors combined with the aggregate statistical biases present in the proposals. For `GBMCMC` the existing proposals introduce some blurring of the priors, as the algorithm explores. The result is that when the algorithm is run in this detailed balance testing mode, the posterior only retains some shape of the prior; however, as a cross-check we should expect some agreement with the prior for a properly functioning search.

An example of this effect is visible in figure 6.1, showing the prior and posterior for sky location when `GBMCMC` is run without any of the new code introduced in this work. The yellow and blue region in the upper part of figure 6.1 comes from the galaxy model. On the right is the prior for reference with the posteriors on the left. The lower half of figure 6.1 shows the same plots, but with the distance prior code enabled. The distance prior is far less efficient

[1]The `--prior` command line flag enables this mode.

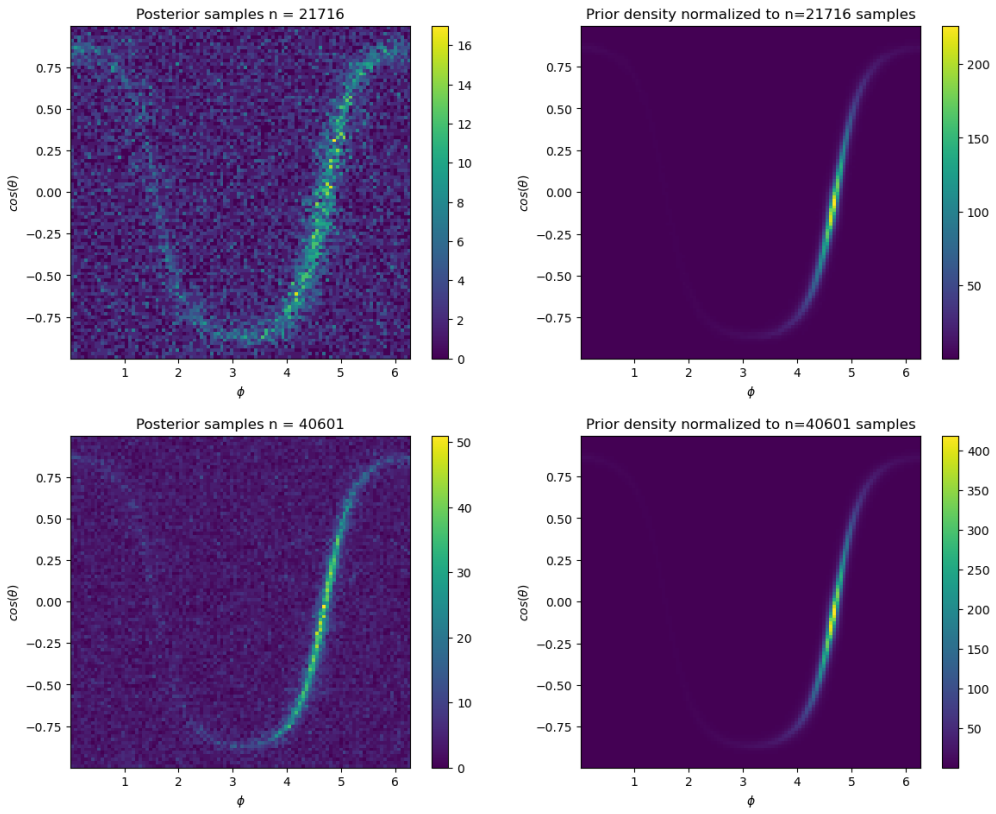

Figure 6.1: Sky location posterior and prior in sky prior mode (top) and distance prior mode (bottom).

at sampling when likelihood is removed from the MCMC algorithm, so the lower graph in 6.1 took ~ 20 times as many MCMC iterations to produce as the upper graph. This inefficiency is less prominent when likelihood is part of the search, mostly because without likelihood the distance prior algorithm has an additional dimension of search space, and therefore a greater volume to explore overall. The thinner posterior distribution in figure 6.1 is also explained by this same effect. The MCMC proposals have freedom to blur the galaxy distribution along distance as well as sky location, so sky location posteriors appear more coherent.

When run with the sky location prior, `GBMCMC` originally drew source amplitude by using an SNR prior combined with the noise model[14]. In volume prior mode, the draw described in section 4.1 is used instead of the SNR prior. Figure 6.2 shows the difference in the

posteriors due to this change. In sky prior mode, the shape of the posterior substantially matches the SNR prior. In volume prior mode, the posterior distribution does not match the SNR prior, because the SNR prior is not in use.

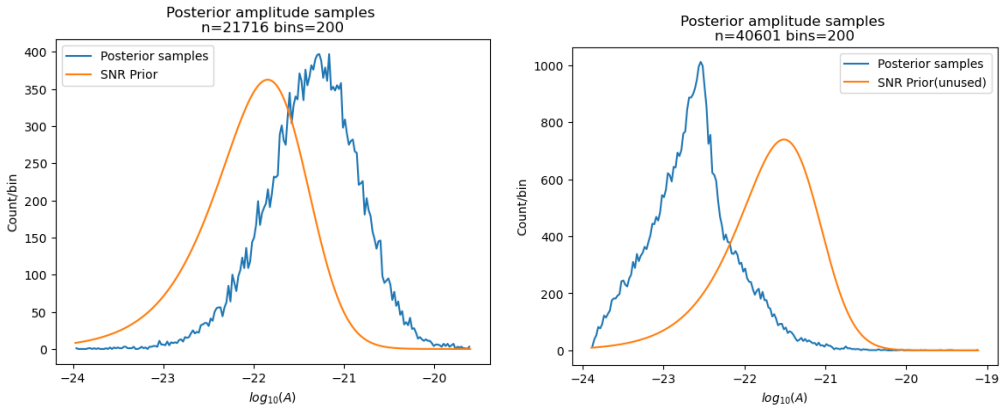

Figure 6.2: Amplitude posterior for sky prior mode (left) and volume prior mode (right) compared to amplitude from SNR prior. SNR prior has been shifted to account for the `GBMCMC` noise model during the run.

In figure 6.2 the tendency for `GBMCMC` in volume prior mode to search a larger search space is also visible in the peak location of the posterior. The peak at lower amplitude in volume prior mode is a result of the distance proposal spending more time exploring the outer volume of its bounding sphere. The most common source accepted is lower amplitude because it is farther away.

Posteriors in distance can also be compared to the distance prior. Figure 6.3 has a comparison for all posterior points across the sky. There is a similar flattening out of the spike in the distance prior as occurs in the sky location histograms. The posterior is also influenced by the uniform nature of the distance proposal function in the tail. The uniform volume density drawn on figure 6.3 is normalized to be comparable to both the prior and posterior to illustrate this influence visually.

The distance component of the galaxy prior has different shapes in different sky directions. Figure 6.4 integrates the posteriors over different sky directions (highlighted in red) to show

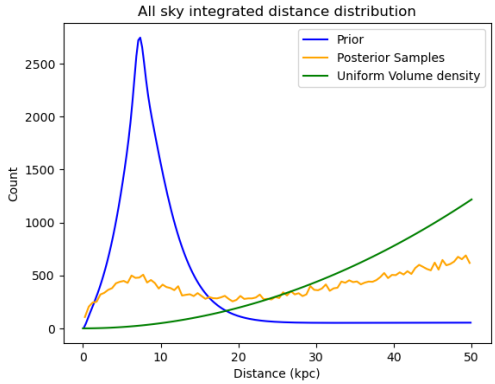

Figure 6.3: Distance posterior and prior in distance prior mode.

the relative agreement between the galaxy prior and posterior samples.

Figures 6.3 and 6.4 show some additional important characteristics of the volume proposal. At distance $D_L = 0$kpc there are consistently posterior samples. Without the contribution of the $g(\mathbf{r})$ portion of the draw function, which is uniform in coordinate r rather than volume, there are not posterior samples at 0 kpc. Additionally, there are posterior samples all the way out to the edge of the bounding sphere. This effect is due to a combination of the isotropic portion of the volume prior and the uniform in volume piece of the proposal distribution.

6.2 Convergence

`GBMCMC` uses an adaptive scheme to gauge the length of the burn-in process of the MCMC search. In summary, sufficiently large changes to the likelihood result in the program restarting from the beginning. This process is managed with an iteration counter. For an MCMC run with a target number of iterations n, the counter starts at $-n$ and counts up. If likelihood changes significantly while the counter is negative, the counter is reset to $-n$. Only posterior samples taken when the counter is greater than zero count toward the catalog that `GBMCMC` produces. This restart behavior is visible in the iteration count graphs of figure 6.5.

With the distance prior enabled, `GBMCMC` takes longer to converge, usually requiring $n =$

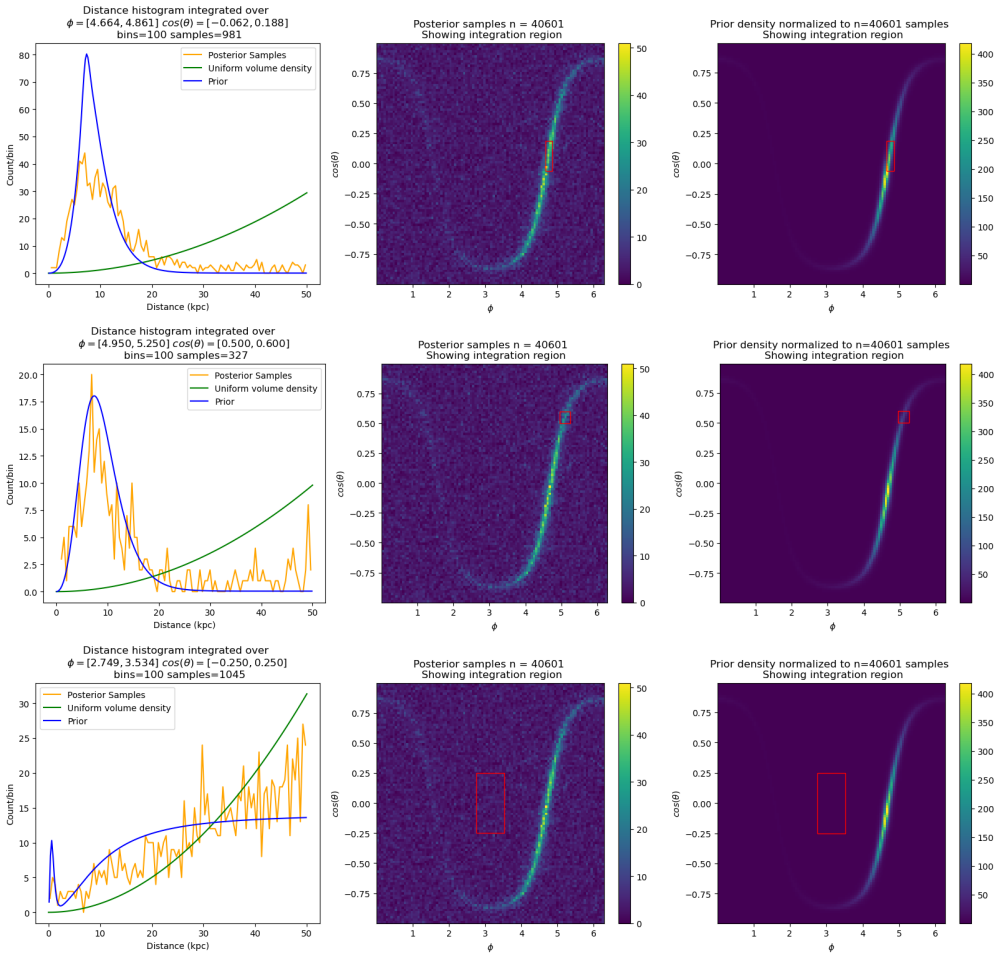

Figure 6.4: Distance posterior and prior in distance prior mode integrated over various regions.

200000 rather than the $n = 100000$ in sky prior mode. Figure 6.5 has a visualization of this effect. In the run shown on the left there is an uptick in the number of UCB sources found and the likelihood near iteration 350000, where the iteration index is greater than zero. With a larger value of n this is solved as shown on the right. Given the higher dimensional and degenerate search space in distance prior mode compared to sky prior mode, a greater number of MCMC iterations are expected for convergence.

While this comparison to the sky prior mode and the stability of source numbers and

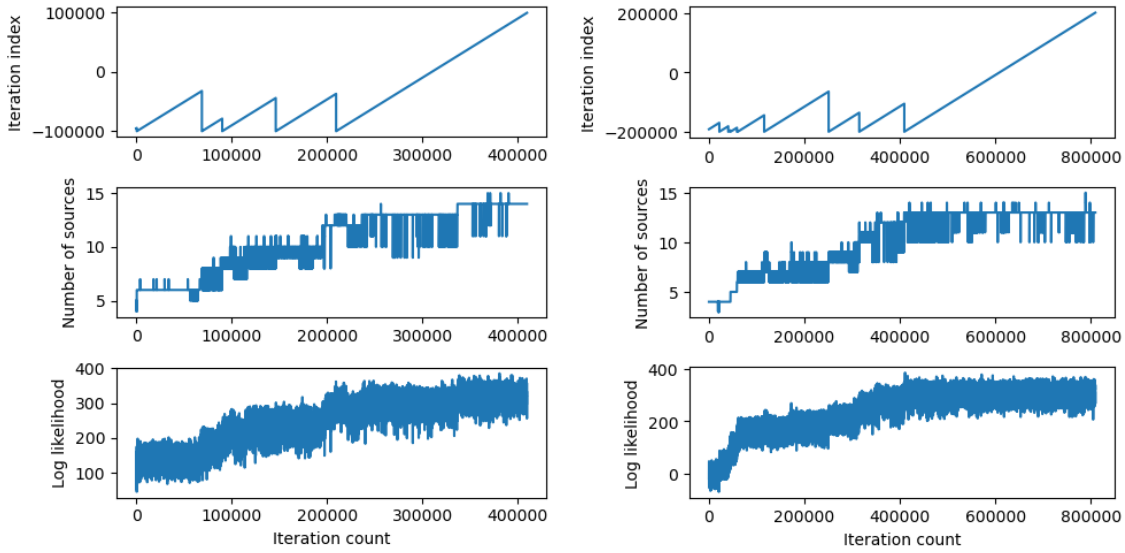

Figure 6.5: Graphs of iteration index, source count, and log likelihood for `GBMCMC` in distance prior mode. A failure to converge ($n = 1 \times 10^5$) is on the left and convergence is shown on the right. ($n = 2 \times 10^5$)

likelihood values provide some evidence of convergence, they do not clearly provide evidence that `GBMCMC` is converging on the correct target function. This is difficult given that the target function for a gravitational wave search is unknown. The following comparisons with simulated data address this difficulty somewhat; however, future work could include more rigorous tests of the convergence of this distance-prior-based MCMC scheme, similar to those employed in [11].

6.3 Detection

The posterior chains from `GBMCMC` must be sent through post-processing in order to create a catalog of sources. The exact procedure is described in [14], and was not altered for this work. The catalog code uses the posteriors from `GBMCMC` as well as the source count appearing most often in the MCMC search to produce posterior distributions for discrete sources. The catalog code is also capable of amending a catalog and keeping a history, but this functionality was not used for this work[14].

Each posterior sample in an RJMCMC chain contains parameters for multiple UCB sources. There is no mapping between posterior samples that would reveal if a source parameter set in each posterior sample corresponded to the same underlying UCB source in the gravitational wave data. It is therefore necessary to have a procedure to label the sources found in the underlying data and assign source parameters from the posterior chains to each found source. Sources found by `GBMCMC`, and posterior distributions aligned with individual sources presented in this section have all been generated using this catalog program from raw `GBMCMC` posteriors.

6.3.1 Hyper-Parameters

The distance prior and related proposal changes each introduced hyper-parameters to the MCMC search. The proposal introduced a weight for the uniform-in-r-coordinate $g(\mathbf{r})$ portion of the draw, and the distance prior introduced a smoothing distance of $a = 10$kpc and a weight for the isotropic contribution to the prior density. The weights for the prior and proposal were chosen by examining the number of sources found by `GBMCMC` over a crowded frequency band around[2] 4mHz. In this band of the LDC Radler data, `GBMCMC` run in sky prior mode finds 14 sources.

Table 6.1: Sources found in distance prior mode by hyper-parameter choice.

Isotropic prior weight	0.0	0.0	0.0	0.0	0.1	0.1	**0.1**	0.1
Uniform r-coordinate proposal weight	0.0	0.1	0.2	0.3	0.0	0.1	**0.2**	0.3
Sources detected	1	6	6	6	1	6	**13**	13

Table 6.1 shows the number of sources found in distance prior mode for each choice of hyper-parameters. The smoothing distance on the prior isotropic contribution was consistently set to 10kpc. The bold column indicates the value of the hyper-parameters in use

[2] 3.9936mHz to 4.0018mHz

for the other results presented here. For the uniform r-coordinate proposal weight 0.2 was chosen over 0.3 to minimize any bias introduced to the proposal distribution by setting a higher value.

6.3.2 UCB Sources

Figure 6.6 shows `GBMCMC` posteriors by source. Analysis was run over the same crowded 4 mHz band with LDC Radler data as input. The parameters of the simulated LDC sources are marked with a red x for a detached binary source, and a red circle for an interacting binary source. Figures 6.7 and 6.8 show corner plots of posteriors for a bright and a dim source in this same band.

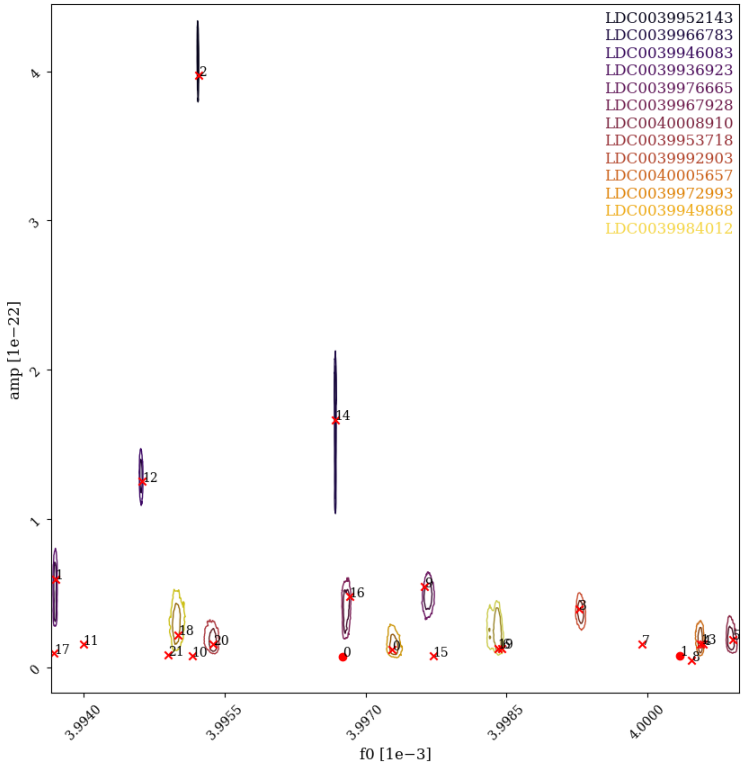

Figure 6.6: Amplitude versus frequency diagram showing injected sources and posteriors in distance prior mode.

`GBMCMC` is good at discerning $f_0, \mathcal{A}, \theta, \phi$. For these parameters, the posteriors have the multivariate Gaussian character we would expect for the bright source, and they are somewhat impacted by instrumental and confusion noise for the dimmer source.

The bright source in figure 6.7 shows very clearly the degeneracy between chirp mass and distance. The simulated bright source is relatively near earth in the galactic plane, but not in the direction of the galactic center. Though the source is quite close, without a clear constraint on its mass, the most likely distance from the posterior is farther away than the injected source distance. The posterior does constrain the source to being in the galactic disk due to the distance prior. By way of comparison, the dimmer source in figure 6.8 has the most likely distance from the posterior near the injected value; however, the uncertainty is much greater than for a bright source.

The rate of frequency change from non-gravitational wave sources ($\dot{f}_{astro}$) is not well constrained by this method. The prior on this parameter is uniform in the current results, and future work could include adding a physically motivated prior to this parameter. The vast majority of LISA UCB sources are expected to be detached binaries[16], so a prior that prefers $\dot{f}_{astro} = 0$ could potentially improve parameter estimation for bright detached sources, or simply make the search more computationally efficient.

6.4 Frequency Band Comparison

Much of the analysis and development of the distance prior was performed by searching the same crowded 4 mHz band in LDC Radler data. I tested several other bands, in order to ensure that the results in the 4 mHz band were generalizable to other bands. The 3.47 mHz band was chosen because it harbors an extremely bright interacting galactic binary in the simulated data, and the 2.31 mHz and 6.01 mHz bands were chosen at random to provide some balance, and look at performance in regions where the bounding sphere of the search is significantly larger or smaller. Table 6.2 summarizes these results.

Table 6.2: Comparison of sources found across several frequency bands.

Frequency band	4 mHz	3.47 mHz	2.31 mHz	6.01 mHz
Bounding sphere radius	49.11 kpc	35.49 kpc	16.47 kpc	115.87 kpc
MCMC iterations including burn-in[3]	6.02×10^5	4.37×10^5	4.54×10^5	4.12×10^5
Sources detected (sky location)	14	11	3	5
Sources detected (distance)	13	10	3	5

Both the 4 mHz and 3.47 mHz fields are crowded, so it is unsurprising that the sky prior and distance prior modes in `GBMCMC` find different numbers of sources. Across the board these support the notion that the volume prior is detecting sources similarly to the sky prior, while needing more compute time. Comparing the 4 mHz iteration number to the others, the size of the bounding sphere appears to have a smaller impact on the compute resources necessary for convergence than how crowded the frequency band is.

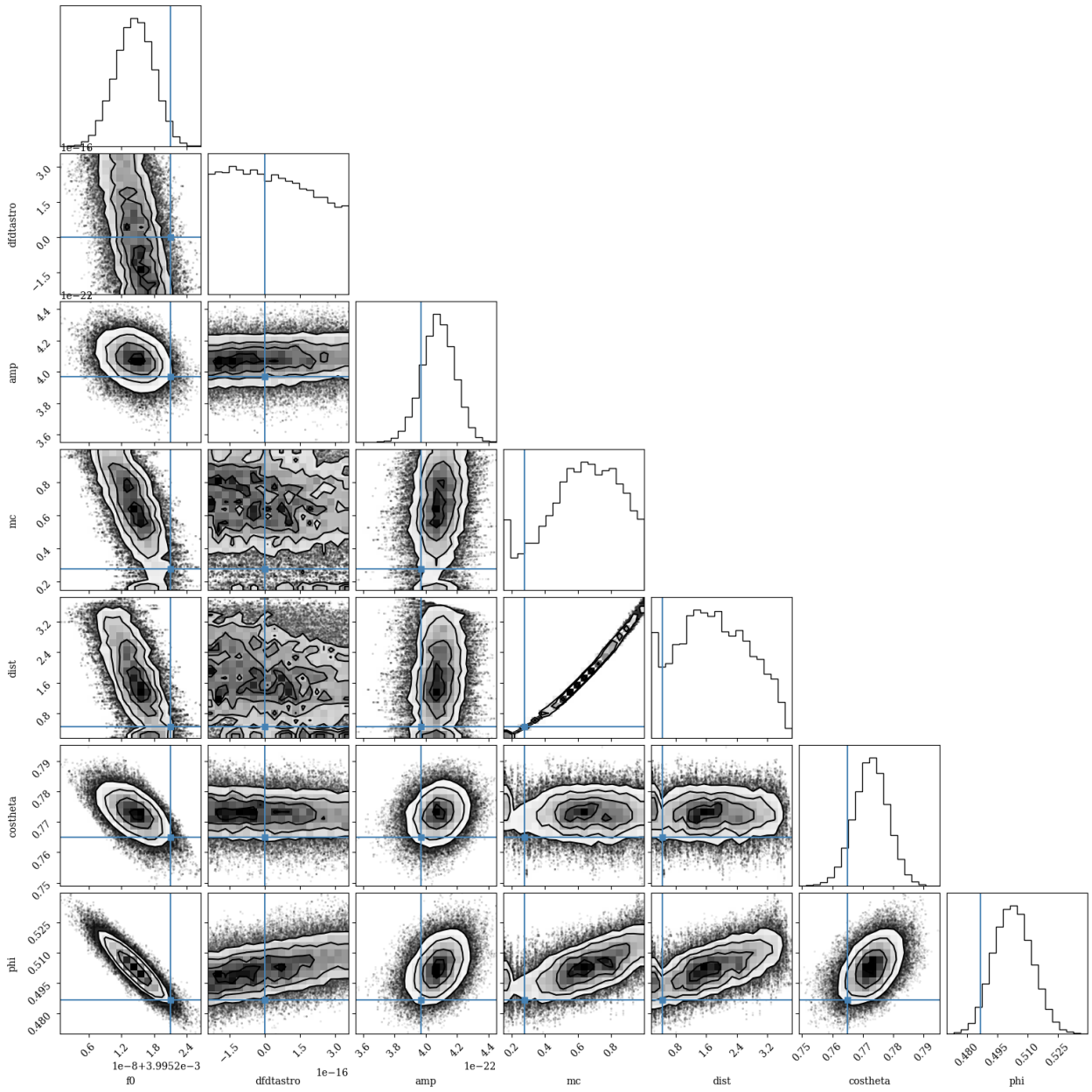

Figure 6.7: Corner plot of distance prior mode posteriors for a bright source in the 4 mHz band (source 2 in figure 6.6).

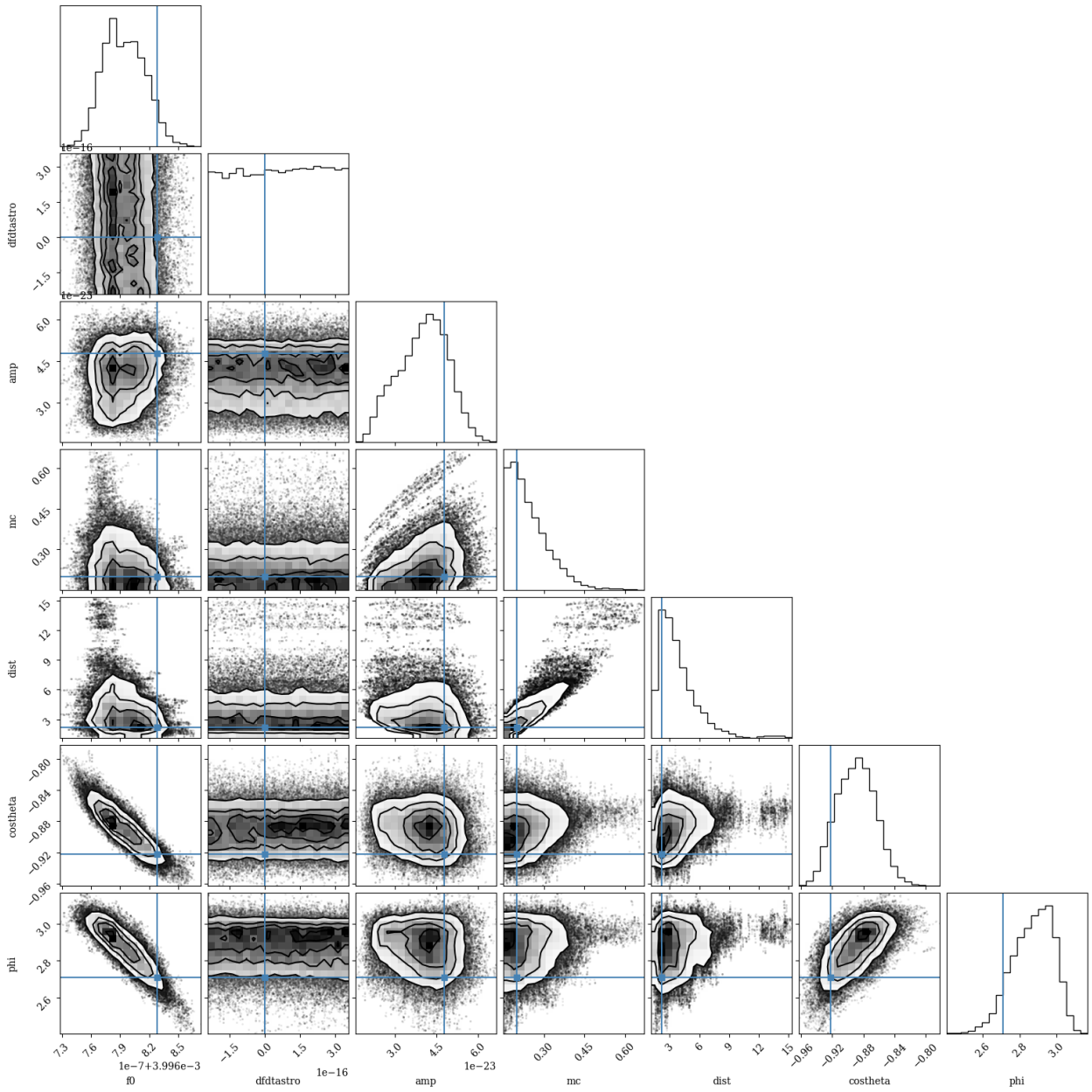

Figure 6.8: Corner plot of distance prior mode posteriors for a dim source in the 4 mHz band (source 16 in figure 6.6).

Chapter 7

CONCLUSION

In pursuit of a solution to the LISA data analysis problem for UCBs, I have modified `GBMCMC` to accommodate a more physically relevant parameter space for incorporating Bayesian priors and producing gravitational wave source catalogs. This better aligns `GBMCMC` with other astrophysical investigations of UCBs and helps pave the way for future multi-messenger observations.

The new prior and proposal added three hyper-parameters to `GBMCMC`. An isotropic prior weight similar to the sky prior (0.1), a uniform-r-coordinate draw weight (0.2), and a smoothing distance (10 kpc). The first two of these parameters have been investigated by comparing sources detected with the sky and distance priors in a 4 mHz band of LDC Radler data. Future work could characterize the third hyper-parameter, or further optimize these hyper-parameters across different bands and simulated data sources.

The necessity of altering the Fisher matrix proposal is also worthy of future investigation, both for its mathematical underpinnings, and how it performs in helping `GBMCMC` leave secondary likelihood maxima. Time[1] did not permit any but the most cursory investigation of this proposal. The slightly sub-par results of the frequency band comparison in section 6.4 could be entirely explained by a subtle inefficiency in this proposal which becomes relevant in crowded frequency bands.

There is almost certainly room for improvement in `GBMCMC`'s estimation of distance by this method. While the degeneracy that D_L participates in looms large, more rigorous tests of convergence, and assessment of `GBMCMC`'s distance estimates over a larger range of frequencies and observation times both hold promise. The introduction of a new parameter space to

[1] ...and my merely rudimentary understanding of information geometry...

GBMCMC opens the door to future investigation of priors on $\dot{f}_{astro}$ and $\mathcal{M}_c$. This may allow GBMCMC to detect interacting binaries, or better constrain distance estimates.

BIBLIOGRAPHY

[1] Quentin Baghi. The LISA Data Challenges. https://arxiv.org/abs/2204.12142v1, April 2022.

[2] John Baker, Jillian Bellovary, Peter L. Bender, Emanuele Berti, Robert Caldwell, Jordan Camp, John W. Conklin, Neil Cornish, Curt Cutler, Ryan DeRosa, Michael Eracleous, Elizabeth C. Ferrara, Samuel Francis, Martin Hewitson, Kelly Holley-Bockelmann, Ann Hornschemeier, Craig Hogan, Brittany Kamai, Bernard J. Kelly, Joey Shapiro Key, Shane L. Larson, Jeff Livas, Sridhar Manthripragada, Kirk McKenzie, Sean T. McWilliams, Guido Mueller, Priyamvada Natarajan, Kenji Numata, Norman Rioux, Shannon R. Sankar, Jeremy Schnittman, David Shoemaker, Deirdre Shoemaker, Jacob Slutsky, Robert Spero, Robin Stebbins, Ira Thorpe, Michele Vallisneri, Brent Ware, Peter Wass, Anthony Yu, and John Ziemer. The Laser Interferometer Space Antenna: Unveiling the Millihertz Gravitational Wave Sky, July 2019.

[3] Jeff Crowder and Neil Cornish. A Solution to the Galactic Foreground Problem for LISA. *Phys. Rev. D*, 75(4):043008, February 2007.

[4] Curt Cutler and Éanna E. Flanagan. Gravitational waves from merging compact binaries: How accurately can one extract the binary's parameters from the inspiral waveform? *Phys. Rev. D*, 49(6):2658–2697, March 1994.

[5] Peter J. Green. Reversible Jump Markov Chain Monte Carlo Computation and Bayesian Model Determination. *Biometrika*, 82(4):711–732, 1995.

[6] Peter J. Green, Nils Lid Hjort, and and Sylvia Richardson. *Highly Structured Stochastic Systems*. Oxford Statistical Science Series. Oxford University Press, Oxford, New York, July 2003.

[7] Piotr Jaranowski, Andrzej Królak, and Bernard F. Schutz. Data analysis of gravitational-wave signals from spinning neutron stars. I. The signal and its detection. *Phys. Rev. D*, 58(6):063001, August 1998.

[8] Kristen Lackeos, Tyson B. Littenberg, Neil J. Cornish, and James I. Thorpe. The LISA Data Challenge Radler Analysis and Time-dependent Ultra-compact Binary Catalogues. *A&A*, 678:A123, October 2023.

[9] Tyson B. Littenberg. A detection pipeline for galactic binaries in LISA data. *Phys. Rev. D*, 84(6):063009, September 2011.

[10] Tyson B. Littenberg. Personal Communication, 2023.

[11] Tyson B. Littenberg and Neil J. Cornish. A Bayesian Approach to the Detection Problem in Gravitational Wave Astronomy. *Phys. Rev. D*, 80(6):063007, September 2009.

[12] Tyson B. Littenberg and Neil J. Cornish. Bayesian inference for spectral estimation of gravitational wave detector noise. *Phys. Rev. D*, 91(8):084034, April 2015.

[13] Tyson B. Littenberg and Neil J. Cornish. Prototype Global Analysis of LISA Data with Multiple Source Types, January 2023.

[14] Tyson B. Littenberg, Neil J. Cornish, Kristen Lackeos, and Travis Robson. Global analysis of the gravitational wave signal from Galactic binaries. *Phys. Rev. D*, 101(12):123021, June 2020.

[15] Michele Maggiore. *Gravitational Waves*. Oxford University Press, Oxford, 2008.

[16] G. Nelemans, L. R. Yungelson, and S. F. Portegies Zwart. The gravitational wave signal from the Galactic disk population of binaries containing two compact objects. *A&A*, 375(3):890–898, September 2001.

[17] G. Nelemans, L. R. Yungelson, M. V. Van Der Sluys, and Christopher A. Tout. The chemical composition of donors in AM CVn stars and ultracompact X-ray binaries: Observational tests of their formation. *Monthly Notices of the Royal Astronomical Society*, 401(2):1347–1359, January 2010.

[18] Frank Nielsen. An elementary introduction to information geometry. *Entropy*, 22(10):1100, September 2020.