DATA ANALYSIS FOR SPACE-BASED GRAVITATIONAL WAVE DETECTORS

by

Jefferson Osborn Crowder

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

of

Doctor of Philosophy

in

Physics

MONTANA STATE UNIVERSITY
Bozeman, MT

April 2006

APPROVAL

of a dissertation submitted by

Jefferson Osborn Crowder

This dissertation has been read by each member of the dissertation committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the Division of Graduate Education.

Neil J. Cornish, Ph. D.

Approved for the Department of Physics

William A. Hiscock

Approved for the Division of Graduate Education

Joseph J. Fedock

STATEMENT OF PERMISSION TO USE

For my father and mother.
Thank you for your encouragemnt
and for your persistance.

ACKNOWLEDGMENTS

Foremost of people to acknowledge for the integral part he played in making this dissertation a reality is my advisor, Dr. Neil Cornish. His ideas and nearly boundless scientific knowledge initiated most of the original work found herein. He may have only been joking the day he asked me if I wanted a job, but it turned out to be the best thing for me. I didn't begin with the idea that I would become a researcher, but with him as an example and by following his guidance I now revel in the process of discovering new facets of gravitational wave astronomy. I cannot thank him enough.

I am indebted to the members of my graduate committee: Dr. William Hiscock for his questions and support which prompted me think deep about many things inside and outside of physics, Dr. Dana Longcope for helping with all the oddball questions I brought to him, Dr. Gregory Francis for showing me you never understand something until you can teach it to someone else, and Dr. Ed Porter for answering my many questions. These men gave of their time and wisdom, I am very grateful.

I would like to thank the current and past members of the Department of Physics: Dr. Louis Rubbo and Dr. Olivier Poujade who were sources of and sounding boards for many of the early ideas that found their way into this work, Lucas Reddinger for his collaboration on genetic algorithms, my wonderful teachers including Drs. George Tuthill, Jeff Adams, and Recep Avci, the excellent staff including Margaret Jarrett, Rose Waldon, Jeannie Gunderson, and Jeremy Gay, and my fellow graduate students who helped me through the years, most recently those who proofread this work Sytil Murphy, Joey Key, Tyson Littenberg, Michael Obland, and Brian Larsen.

Finally, I would like to thank my parents. My mother gave me unending support through my long career as a student. My father was ever persistent telling me that nothing was as important as finishing my education. That I stuck with it and have reached this point is due mostly to him. I know he is smiling down on me today.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CONVENTIONS

Geometric units, $G = c = 1$, are used throughout this dissertation. However, frequencies are given in Hertz. The sign conventions for general relativity used in this dissertation follow those used in *Gravitation* by Misner, Thorne, and Wheeler [1].

Rank-2 tensors are designated by plain bold symbols, $\mathbf{T}$, while four-vectors are denoted by italicized bold symbols, $\boldsymbol{v}$. Purely spatial vectors are designated with a vector symbol over the variable $\vec{x}$.

When using spacetime indices, Greek letters are used to range over the time and spatial components $0, 1, 2, 3$, while Latin letters are used for spatial components only. A bar over the index, or any coordinate label, denotes a change in coordinates,

$$\xi^{\bar{\mu}} = \Lambda^{\bar{\mu}}_{\alpha} \xi^{\alpha}. \tag{1}$$

Partial differentiation follows the standard comma notation,

$$\xi^{\mu}_{,\nu} = \partial_{\nu} \xi^{\mu} = \frac{\partial \xi^{\mu}}{\partial x^{\nu}}. \tag{2}$$

Covariant differentiation is signified with a semicolon,

$$\xi^{\mu}_{;\nu} = \frac{\partial \xi^{\mu}}{\partial x^{\nu}} + \xi^{\alpha} \Gamma^{\mu}_{\alpha\nu}. \tag{3}$$

ABSTRACT

With the launch of the Laser Interferometer Space Antenna (LISA) expected for the next decade, the nascent field of gravitational wave astronomy will be taking a giant leap forward. The data that will be gathered from space-borne gravitational wave detectors such as LISA will provide an expansive look through a new window on the Universe. This dissertation is presented to help open that window by exploring some of the techniques and methods that will be needed to understand the data from these detectors.

The first original work presented here investigates the resolution of LISA and follow-on space-based gravitational wave missions. This work presents the methods of measuring the precision of these detectors and gives results for their resolving power for a large class of expected gravitational wave sources.

The second original investigation involves the effect that multiple gravitational wave sources will have on the resolution of LISA. Previous results concerning detector resolution were limited to isolated sources of gravitational waves. As LISA is an all-sky detector, it is necessary to understand the role played by concurrent detection of numerous sources. This work derives an extension of the Fisher Information Matrix approach for determining parameter resolution, and applies it to multiple sources for LISA.

The next original work is an exploration of the method of genetic algorithms on the problem of extracting the binary parameters of gravitational wave sources from the LISA data stream. These are global algorithms providing a means to cover the entire search space of parameter values. This work describes the basics of and provides the results for such genetic algorithm-based searches, with a focus on improving algorithm efficiency.

The last original work included is a study of Markov Chain Monte Carlo (MCMC) methods applied to parameter extraction of gravitational wave sources in the LISA data stream. This work shows how an MCMC approach provides a global means of both searching for and characterizing the distributions of the source parameters. Results also show that distributions found by this global method match with previous approaches that were limited to regions local to the source parameters.

CHAPTER 1

GRAVITATIONAL WAVES IN GENERAL RELATIVITY

Introduction

In 1864, James Clerk Maxwell presented to the Royal Society of London for the Improvement of Natural Knowledge the complete set of equations that describe classical electromagnetism. This set of equations, which have become known as Maxwell's Equations, were built from the works of previous physicists and mathematicians including Michael Faraday, Karl Friedrich Gauss, Charles Coloumb, and Andrè-Marie Ampère. The genius of Maxwell was to include a new set of terms that corrected an inconsistency in Ampère's Law. This set of terms described the effect of a time-varying electric field on electric currents and magnetic fields in materials. In his search to find a solution to the inconsistency in Ampère's Law, Maxwell found the key that opened the door to a new description of the universe.

A year after his initial presentation, Maxwell discovered that with his new terms in place, he could solve the complete set of equations using a simple sine wave [2]. Even more simple was the fact that the solution was limited to one group velocity for the waves; a velocity that was independent of the person observing the wave. The group velocity Maxwell calculated for these waves was so close to the speed of light it gave him reason to conclude light itself was probably a type of electromagnetic disturbance.

Up until that time all of the types of waves that physicists were familiar with traveled in one type of media or another, for example, sound waves traveled through liquids, gases, and solids. So the natural assumption of the physicists of the time was that light waves must also travel through some type of as yet undetected medium, and the group velocity that Maxwell had calculated was with respect to that medium. The name they gave to the medium was the luminiferous ether.

In 1887 Albert Michelson and Edward Morley attempted to show the existence of the luminiferous ether. In their experiment light was directed along two paths that were at right angles to each other [3]. The expectation was that they would measure different velocities for the two light waves, due to differences in how they moved with respect to the luminiferous ether. Their expectation was not met. Instead, their experiment became one of the most famous null results in the history of physics. The implication was not only that light waves propagated through vacua, but also that the speed with which light propagates through a vacuum is a constant for any and all observers.

In 1905, Albert Einstein used the constancy of the speed of light in vacuum as the first of two postulates he would set forth as the foundation of his theory of Special Relativity [4]. The second of the postulates mandated that the laws of physics had to be the same in any inertial (non-accelerating) reference frame. From these two postulates Special Relativity explained the results of the Michelson-Morley experiment, led to the joining of space and time into one physical construct, spacetime, and showed the relationship between mass and energy. In 1907, while seeking to generalize Special Relativity, Einstein introduced a third postulate, known as the Equivalence Principle, which stated that a uniform acceleration is equivalent to a gravitational field [5]. From this realization, Einstein was able to expand the theory of Special Relativity into the theory of General Relativity [6].

General Relativity is encapsulated in ten independent, non-linear partial differential equations, known as the Einstein Equations, which can be written in a shorthand notation as:

$$\mathbf{G} = 8\pi\mathbf{T}\,, \tag{1.1}$$

where $\mathbf{G}$ is the Einstein tensor describing the curvature of the spacetime, and $\mathbf{T}$ is the stress-energy tensor, describing the distribution of mass and energy in the spacetime. A consequence of the nature of $\mathbf{T}$ as the distribution of mass and energy is that it follows the conservation law,

$$T^{\mu\nu}{}_{;\nu} = 0. \tag{1.2}$$

As these partial differential equations are non-linear, solutions are difficult to determine. In fact, only a handful of physically useful solutions, dealing with idealized situations, have been found to the Einstein Equations and no general solution exists. To simplify the task of understanding General Relativity, and how it predicts the presence of gravitational waves, we will begin by looking at what happens in the theory when we restrict ourselves to the weak-field limit.

Linearized General Relativity

In the weak-field limit of General Relativity, a small metric perturbation, $\mathbf{h}$, is added to a flat spacetime. The smallness of the perturbation is measured relative to the flat spacetime metric of Special Relativity, $\eta_{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1)$.

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \qquad\qquad |h_{\mu\nu}| \ll 1 \tag{1.3}$$

For the weak-field limit an expansion of (1.1) in powers of $h_{\mu\nu}$, the terms that are higher than first-order in $h_{\mu\nu}$ will be negligible. Thus they may be discarded without an appreciable loss of accuracy. This leaves only the terms that are linear in $h_{\mu\nu}$. This formalism is known as “the linearized theory of gravity.” The Einstein tensor to first-order in $h_{\mu\nu}$ is given by:

$$G_{\mu\nu} = -\frac{1}{2}(h_{\mu\nu,\alpha}{}^{,\alpha} + \eta_{\mu\nu}h_{\alpha\beta}{}^{,\alpha\beta} + h^{\alpha}{}_{\alpha,\mu\nu} - h_{\mu\alpha,\nu}{}^{,\alpha} - h_{\nu\alpha,\mu}{}^{,\alpha} - \eta_{\mu\nu}h^{\alpha}{}_{\alpha,\beta}{}^{,\beta}) + \mathcal{O}[h^2]. \quad (1.4)$$

Defining a so-called trace-reversed perturbation as:

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h_{\alpha}{}^{\alpha}\,, \quad (1.5)$$

one can simplify (1.4) to,

$$G_{\mu\nu} = -\frac{1}{2}(\bar{h}_{\mu\nu,\alpha}{}^{,\alpha} + \eta_{\mu\nu}\bar{h}_{\alpha\beta}{}^{,\alpha\beta} - \bar{h}_{\mu\alpha,\nu}{}^{,\alpha} - \bar{h}_{\nu\alpha,\mu}{}^{,\alpha}) + \mathcal{O}[\bar{h}^2]. \quad (1.6)$$

This is further simplified if $\bar{h}_{\alpha\beta}{}^{,\beta} = 0$, since only the first linear term in (1.6) would remain. However, to impose these four conditions at least one gauge must exist in which these conditions are true. A class of such gauges does exist and these gauges are known collectively, and somewhat ungrammatically, as the Lorentz gauge. The condition for a gauge transformation,

$$h_{\mu\nu} \rightarrow h_{\mu\nu} - \xi_{\mu,\nu} - \xi_{\nu,\mu}\,, \quad (1.7)$$

to be a member of the Lorentz gauge is that ξ_{μ} is a solution to the source-free wave equation:

$$\xi_{\mu\nu,\alpha}{}^{,\alpha} = 0. \quad (1.8)$$

Thus the field equations for linearized General Relativity in the Lorentz gauge are simply:

$$\bar{h}^{\mu\nu,\alpha}_{,\alpha} = -16\pi \mathrm{T}^{\mu\nu}. \quad (1.9)$$

Gravitational Waves

Clearly, when $T^{\mu\nu} = 0$ the field equations shown in (1.9) reduce to the source-free wave equation. Just as Maxwell's Equations admit wave solutions in a vacuum for light waves in electromagnetism, the Einstein Equations admit wave solutions in a vacuum for gravitational waves in linearized General Relativity. One class of solutions to these equations are monochromatic plane waves:

$$\bar{h}^{\mu\nu} = \Re\left[A^{\mu\nu}\exp\left(ik_\alpha x^\alpha\right)\right]. \tag{1.10}$$

Enforcing the Lorentz gauge on the first derivative of (1.10) shows that $A^{\mu\nu}$ is orthogonal to $\mathbf{k}$, $A^{\mu\nu}k_\nu = 0$. These orthogonality constraints take up four degrees of freedom for the wave. Another four will be taken up when a specific gauge is chosen (as shown below), leaving $A^{\mu\nu}$ with just two degrees of freedom out of the starting ten that show up in the Einstein Equations.

A useful choice of gauge to show how these two degrees of freedom manifest themselves in gravitational waves is known as the transverse-traceless (TT) gauge. The transverse nature of the gauge is set by the requirement that $A_{\mu\nu}u^\nu = 0$, where $\mathbf{u}$ is a 4-velocity of some general observer. The gauge is traceless in that it also requires $A_\mu{}^\mu = 0$. A quick glance at (1.5) shows that the traceless condition mandates $\bar{h}_{\mu\nu} = h_{\mu\nu}$.

Explicitly applying these conditions to a gravitational wave allows one to see the nature of the gravitational wave in the TT gauge. For someone at rest with the general observer ($\mathbf{u} \rightarrow (1,0,0,0)$), the transverse conditions leads to the condition on the gravitational wave, $h_{\mu 0} = 0$, since $A_{\mu 0}u^0 = 0$. This information used in conjunction with the orthogonality of the gravitational wave with the wave vector leads to the condition, $h_{ij,j} = 0$ (summing over j). If one assumes a gravitational wave traveling along the positive z-axis, $h_{\mu z} = 0$, since $\mathbf{k} \rightarrow (\omega, 0, 0, k)$. Lastly, applying the traceless condition gives $h_{xx} = -h_{yy}$. Thus, a gravitational wave in the TT gauge traveling along the z-axis has an amplitude tensor given by:

$$A_{\mu\nu}^{TT} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & A_{xx} & A_{xy} & 0 \\ 0 & A_{xy} & -A_{xx} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{1.11}$$

This representation can be decomposed into two distinct tensors, one depending only on A_{xx}, and the other only depending on A_{xy}. These represent the two polarizations of the gravitational wave. Defining $A_+ \equiv A_{xx}$ and $A_\times \equiv A_{xy}$, and the polarization tensors $\mathbf{e}^+ = \hat{x} \otimes \hat{x} - \hat{y} \otimes \hat{y}$ and $\mathbf{e}^\times = \hat{x} \otimes \hat{y} + \hat{y} \otimes \hat{x}$, the amplitude tensor is written as:

$$\mathbf{A}^{\mathrm{TT}} = A_+ \mathbf{e}^+ + A_\times \mathbf{e}^\times. \tag{1.12}$$

Substituting this into Equation (1.10), we arrive at an expression for the solution to gravitational waves in the TT gauge in the weak field limit:

$$\bar{\mathbf{h}}^{TT} = \Re \left[A_+ \mathbf{e}^+ + A_\times \mathbf{e}^\times \right] \exp\left(i\mathbf{k} \cdot \mathbf{x} \right). \tag{1.13}$$

Thus the effect caused by introducing this perturbation, $\mathbf{h}$, is an oscillatory disruption of the background metric. This disruption can be detected by observing how the oscillation of the metric affects distances between freely moving objects in the spacetime. Figure 1.1 [7] shows the responses of a ring of test masses to the two polarizations ($+$ and $\times$) of a gravitational wave passing though the ring in the positive z-direction. These motions are a measurable effect that will allow us to detect the presence of gravitational waves. Details of some of the past, present, and future detectors are included in the Detecting Gravitational Waves section of this chapter and in Chapter 2.

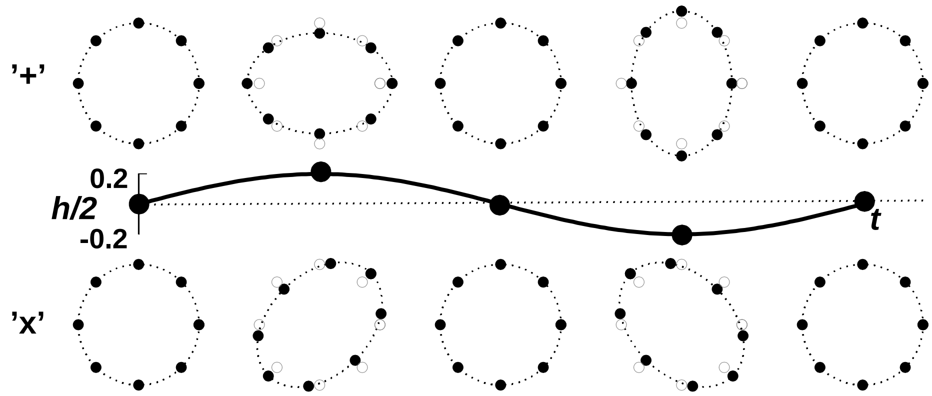

Figure 1.1: The responses of a ring of test masses as a gravitational wave passes through the ring traveling in the positive z-direction. The top row shows the effect of a gravitational wave with a + polarization. The bottom row shows the effect of a gravitational wave with a × polarization. The effects of the two waves are similar, but rotated 45° from each other, caused by the spin-2 nature of the graviton.

The superposition of plane waves of the type shown in Equation (1.13), without being restricted to travel in the positive z-direction, is the most general solution to the source-free wave equation:

$$\bar{h}^{TT}_{\mu\nu} = \Re\left[\int_{\infty}^{\infty} A_{\mu\nu} \exp{(ik_\alpha x^\alpha)} d\vec{k}\right]. \tag{1.14}$$

Generation of Gravitational Waves by Binary Systems

As this work focuses on the gravitational waves emanating from binary systems, preference will be given here to the generation of gravitational waves from such systems. The remaining question regards the nature of the amplitudes A_+ and $A_\times$ appearing in $A_{\mu\nu}$ of Equation (1.14). To understand the form of these amplitudes one looks to an approximate solution to Equation (1.9) with a simplified source term. The assumptions made here are that the source is slow-moving, localized, and oscillating at an angular frequency, ω.

$$T_{\mu\nu} = \Xi_{\mu\nu} \exp\left(-i\omega t\right). \tag{1.15}$$

Solutions exist for the gravitational wave produced by this type of source, and are of the form:

$$\bar{h}_{\mu\nu} = S_{\mu\nu} \exp\left(-i\omega t\right). \tag{1.16}$$

As one moves away from the location of the source, $S_{\mu\nu}$ asymptotically takes the form of a combination of inward and outward traveling spherical waves:

$$S_{\mu\nu} = \frac{A_{\mu\nu}}{r} \exp\left(i\omega t\right) + \frac{B_{\mu\nu}}{r} \exp\left(-i\omega t\right). \tag{1.17}$$

Being interested in the gravitational waves emitted by the source, the term describing inward traveling waves is set to zero ($B_{\mu\nu} = 0$).

With this solution, integrating both sides of Equation (1.9) over the volume of a sphere completely containing the source, neglecting terms of order r^{-2} and r^{-1} terms which are second order in the light travel time across the radius of the source volume, leads to

$$Z^{\mu\nu} = 4 \exp\left(i\omega t\right) \int T^{\mu\nu}\, d^3x\,, \tag{1.18}$$

where $Z^{\mu\nu} = A^{\mu\nu} \exp\left(i\omega t\right)$. One can use the localized nature of the source in conjunction with Equation (1.2) and Gauss' Law to arrive at,

$$\omega t Z^{\mu 0} = 4i \exp{(i\omega t)} \oint T^{\mu k} n_k \, dS \,, \tag{1.19}$$

where the surface for the integral is outside of the source, so the stress-energy tensor is zero. This leads to the conclusion that $Z^{\mu 0} = A^{\mu 0} = 0$.

Using the slow motion assumption, $T^{00} \approx \rho$ (the Newtonian mass density), with the virial theorem,

$$\frac{\partial^2}{\partial t^2} \int T^{00} x^i x^j \, d^3x = 2 \int T^{ij} \, d^3x \,, \tag{1.20}$$

one identifies the left hand side of the equation as the second derivative with respect to time of the quadrupole moment tensor of the source, $I^{ij} \equiv \int \rho x^i x^j d^3x$. Thus one can write the solution for the generated gravitational wave as:

$$\bar{h}_{ij} = -2\omega^2 I_{ij}(t) \exp{(i\omega r)}/r. \tag{1.21}$$

Transforming to the TT gauge this solution is:

$$\bar{h}^{TT}_{xx} = -\bar{h}^{TT}_{yy} = -\omega^2 (\mathcal{I}_{xx} - \mathcal{I}_{yy}) \exp{(i\omega r)}/r \,, \tag{1.22}$$

$$\bar{h}^{TT}_{xy} = -2\omega^2 (\mathcal{I}_{xy}) \exp{(i\omega r)}/r \,, \tag{1.23}$$

$$\bar{h}^{TT}_{zi} = 0. \tag{1.24}$$

where $\mathcal{I}_{ij} \equiv I_{ij} - \frac{1}{3}\delta_{ij} I^k{}_k$ is the reduced quadrupole moment tensor.

In the case of a system of two stars, separated from each other and spherical, with masses M_1 and M_2, orbiting the center of mass of the system with orbital frequency Ω, at distances R_1 and R_2 in the xy plane, the quadrupole moment tensor is given by:

$$I_{xx} = \frac{1}{2}(M_1 R_1^2 + M_2 R_2^2)(1 + \cos(2\Omega t + \varphi_o))\,, \tag{1.25}$$

$$I_{yy} = \frac{1}{2}(M_1 R_1^2 + M_2 R_2^2)(1 - \cos(2\Omega t + \varphi_o))\,, \tag{1.26}$$

$$I_{xy} = \frac{1}{2}(M_1 R_1^2 + M_2 R_2^2)\sin(2\Omega t + \varphi_o)\,, \tag{1.27}$$

$$I_{zi} = 0, \tag{1.28}$$

where φ_o sets the initial azimuthal position of the system with respect to the positive x-axis. This leads to a solution for those gravitational waves propagating perpendicular to the orbital plane (assumed here to be the positive z-direction):

$$\bar{h}_{xx} = -\bar{h}_{yy} = -\frac{4M_1M_2}{rR}\cos(2\Omega(t-r) + \varphi_o)\,, \tag{1.29}$$

$$\bar{h}_{xy} = -\frac{4M_1M_2}{rR}\sin(2\Omega(t-r) + \varphi_o)\,, \tag{1.30}$$

$$\bar{h}_{zi} = 0\,, \tag{1.31}$$

where R is the binary separation distance given by Kepler's Third Law, $R^3\Omega^2 = (M_1 + M_2)$.

These results may be generalized to gravitational waves propagating in the $\hat{k}$-direction by introducing an orthonormal triad $\{\hat{p}, \hat{q}, \hat{k}\}$, in which $\hat{p}$ is perpendicular to $\hat{k}$, and $\hat{q} = \hat{p} \times \hat{k}$. Astronomers define the angle of inclination of the orbital plane of the binary system, ι, as the angle between the angular momentum vector of the binary system and the line of sight of an observer to the center of mass of the binary system. Applying this definition and orthonormal basis the non-zero components of the gravitational wave tensor in the TT gauge are given by:

$$\bar{h}^{TT}_{pp} = -\bar{h}^{TT}_{qq} = \frac{2M_1M_2}{rR}(1 + \cos^2(\iota))\cos(2\Omega(t-r) + \varphi_o)\,, \tag{1.32}$$

$$\bar{h}_{pq}^{TT} = \frac{4M_1M_2}{rR}\cos(\iota)\sin(2\Omega(t-r)+\varphi_o). \tag{1.33}$$

With this general expression one can determine the effect of gravitational waves on a detector built either on the Earth, or in space.

Detecting Gravitational Waves

In Figure 1.1 the reaction of test masses to gravitational waves was shown. Looking at the effect of the masses, one can imagine a simple description of a gravitational wave detector. By measuring the positions of a grouping of freely moving masses, one can look for the changes in distance between the masses that are predicted by theory. Detection of a gravitational waves will not simply provide a theoretical boon, supporting the theory of General Relativity, but it will also provide us with information about the sources of the waves. Thus, operational gravitational wave detectors will open up the new field of Gravitational Wave Astronomy. What is simple in concept, however, is sometimes not as simple to design and build. The remainder of this chapter will focus on the broad features of the design of three classes of gravitational wave detectors.

The detectors that most intuitively follow from imaging the motions of freely moving masses under the influence of a gravitational wave are the currently proposed space-based detectors. The Laser Interferometer Space Antenna (LISA) [8] is made of three freely moving spacecraft orbiting the sun in a triangular arrangement whose sides are 5×10^9 meters long. Each spacecraft is a protective shell for two proof masses which are falling on geodesics around the sun. The spacecraft also contain a laser system which will be used to determine changes in the positions of the proof masses as a function of time. The LISA mission will focus on detecting gravitational waves from supermassive black hole (SMBH) binary systems merging, extreme mass ratio inspirals (EMRIs), and galactic binary systems. Chapter 2 will give a more detailed treatment of space-based gravitational wave detectors, including noise sources and their effects.

One class of ground-based detectors, which are similar to the space-based detectors, are the detectors which also use interferometry. Detectors like the Laser Interferometer Gravitational wave Observatory (LIGO) in the United States, the Gravitational wave Earth Observatory 600 (GEO600) in Germany, VIRGO in Italy, and the Tokyo Advanced Mediumscale Antenna 300 (TAMA300) in Japan use lasers to precisely measure the changes in the optical path length of two orthogonal cavities oriented, similar to those used in the Michelson-Morley experiment, in the shape of an 'L.' These detectors are searching for gravitational waves from the last stages of the inspiral for stellar remnant binary systems, distorted spinning neutron stars, supernovae, and cosmic string cusps. For ground-based detectors gravitational gradients and seismic events in the Earth are a large factor at lower frequencies (< 10 mHz) making detection at such frequencies impossible. Currently LIGO, with its two facilities located near Livingston, Louisiana and Hanford, Washington (shown in Figure 1.2), is the only detector operating at design sensitivity. Providing that theoretical predictions of General Relativity are correct, detection of a gravitational wave by LIGO will depend on the rate at which inspirals occur in the volume of space close enough to Earth that its signal can be found amidst the noise. Advanced LIGO is a follow-on project to the LIGO detector which is expected to exceed LIGO's sensitivity by about a factor of fifteen. This improvement will lead to a volume coverage of the advanced detector that is over $3,000$ times greater than the current LIGO. Also, the frequency range is expanded for Advanced LIGO, which will be sensitive to gravitational waves between ~ 10 Hz and ~ 2.5 kHz.

The other prevalent class of ground-based detectors are the resonant mass or "bar detectors." The names come from fact that these detectors should resonate at

Figure 1.2: Aerial photographs of the LIGO facilities in Washington (top) and Louisiana (bottom).

certain, narrow frequency bands when gravitational waves pass through them and many are bar-like in shape (though spherical “bar detector” designs are being investigated). Current detectors include the Gravitational Radiation Antenna In Leiden (GRAIL) located in the Netherlands (with a second MiniGRAIL being constructed in Brazil), ALLEGRO in the United States, the Antenna Ultracriogenica Risonante per l’Indagine Gravitazionale Astronomica (AURIGA) and the the Nautilus and Exporer detectors of the Rome One Group (ROG) in Italy. The galactic sources they are searching for include gravitational collapse of stars, merger and ringdown of compact objects with masses $\sim 10M_{\odot}$, and non-axisymmetric instabilities of spinning compact objects. Resonant mass detectors are super-cooled to limit thermal noise within the mass. Acoustic and seismic noise also place limits on the sensitivities of these detectors.

CHAPTER 2

SPACE-BASED GRAVITATIONAL WAVE DETECTORS

Introduction

The first gravitational wave detectors were conceived and built by Joseph Weber in the late 1950s and 1960s. Weber's resonant mass detectors used aluminum cylinders suspended by wires from the ceiling of his laboratory at the University of Maryland, and also at the Argonne National Laboratory in Illinois. Weber's attempts to measure gravitational waves [9, 10] did not produce results that were accepted by the scientific community, but his ideas laid the groundwork for the gravitational wave detectors that would follow.

One of Weber's ideas was to use an interferometer, not unlike the one used by Michelson and Morley in their famous experiment, to measure the distance along two paths. This idea was furthered by Weiss [11] and Forward [12]. This idea has now been implemented in ground-based observatories around the world. At the present time there are four such detectors: LIGO, with is two facilities located near Livingston, Louisiana and Hanford, Washington; GEO 600 located near Hannover, Germany; VIRGO near Pisa, Italy; and TAMA 300, near Tokyo, Japan. These detectors are each searching for small variations in the length along their two orthogonal paths, which are laid out in the shape of an 'L'. They are searching for gravitational wave sources with frequencies between roughly 10 and $10,000$ Hz. This range is sometimes called the "high frequency range." In the high frequency range, gravitational wave sources are not long-lived, and astrophysical estimates for the rates of occurrence vary by orders of magnitude. The current best estimates suggest that positive detection for these detectors will be difficult for observation times even as long as a decade, with the detectors operating at design sensitivity. Lower frequencies are not detectable by ground-based observatories due to the gravitational and vibrational disturbances present in and on the Earth. This limitation has led the gravitational wave community to look to space.

LISA is a proposed space-based gravitational wave detector. It will be sensitive to frequencies in the range $\sim 10^{-4}$ to ~ 1 Hz. In this "low frequency range" there will be many thousands of sources, which will be emitting gravitational waves continuously, above the design sensitivity level of LISA, providing almost immediate detection of gravitational waves when LISA is operational.

The LISA observatory employs three free-flying spacecraft that will form an interferometer in space (see Figure 2.1, taken from [13]). Their orbital paths will keep them located at the vertices of a nearly equilateral triangle, with sides 5×10^9 meters long. The guiding center of LISA will follow the same orbit around the Sun as the Earth (R = 1 AU), trailing the Earth by 20°. The plane containing the three spacecraft will be inclined 60° with respect to the ecliptic. Figure 2.2, taken from [7] shows the orbital path of LISA around the Sun. Each of the three spacecraft will contain two lasers directed toward the other two spacecraft. The lasers will provide a means to measure the small changes in distance along each of the three arms of the detector.

Detector Operation

Once LISA is operational, it will act as an all-sky antenna. Though Figure 2.3, taken from [14] shows that LISA is more sensitive to sources located in directions perpendicular to the plane of the detector, gravitational waves coming from any direction will affect the antenna. Given the free-fall requirements of the orbital path for LISA, one cannot point the antenna as can be done with a telescope. Gravitational

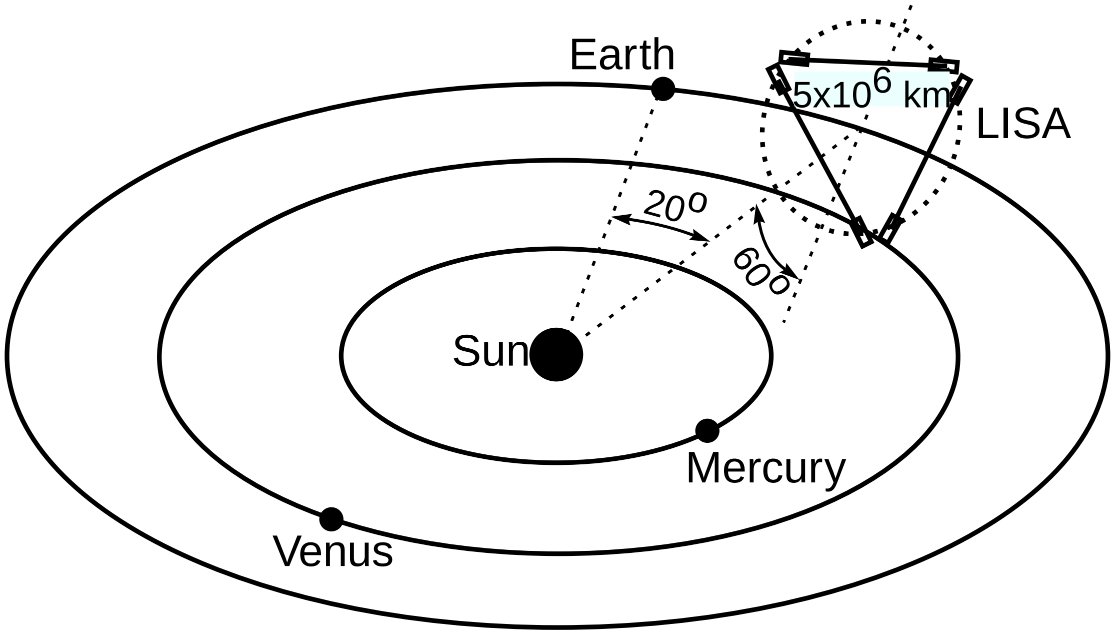

Figure 2.1: Orbital configuration of LISA.

waves from any source in the Universe will affect the motion of the LISA proof masses as they pass through the detector, but most of these waves will be too weak to detect above the inherent noise of the detector. There will still be many thousands of sources whose signals will be detectable above the noise with proper data analysis techniques. The ability to do gravitational wave astronomy with LISA depends on ability to do four things. First is the ability to model the gravitational waveforms of the various sources of gravitational waves. Second is the ability to include the motions and sensitivities of LISA with those models so that the theoretical data will closely resemble the true LISA responses. Third is the ability of LISA to accurately measure the small changes in spacecraft separation induced by the interaction with the actual waveforms. Fourth is the ability to perform the necessary data analysis, working backwards from the LISA data, through the models of detector orbit and waveform generation to determine the parameters of the myriad gravitational wave sources.

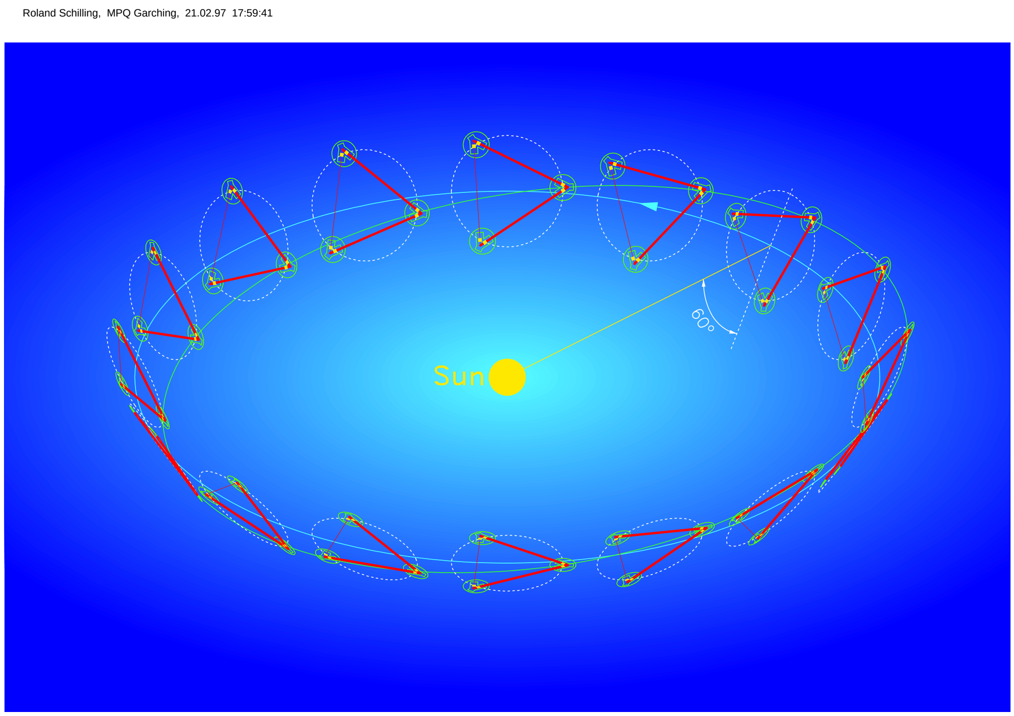

Figure 2.2: Orbital path of LISA around the sun.

Interferometry with LISA

Inside each of the three spacecraft are two proof masses, which will be falling freely along their geodesics around the Sun during operation. One of the purposes of the spacecraft is to protect the proof masses from external perturbations (e.g. the solar wind). A second purpose of the spacecraft is to provide platforms for the laser systems. The proof masses are the reference points for the laser systems from which phase information will be taken. By noting the changes in the phase, one can determine the change in distance traveled by the laser photons to exquisite precision and accuracy. With LISA the distances across which the light is traveling is great (5×10^9 m), allowing for ultra-precise measurements of the strain (defined as the change in length of the arm divided by its overall length, strain $= \Delta L/L$) induced

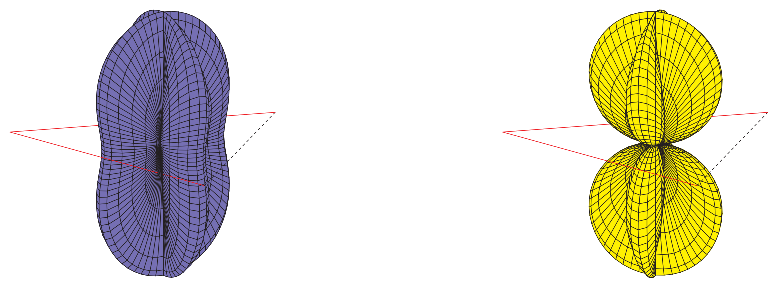

Figure 2.3: Representations of the detector beam patters F^{+} and $F^{\times}$ in the low frequency limit.

by the passage of a gravitational wave. The goal for LISA is to measure the strain to $\sim 10^{-21}$, which corresponds to arm length changes of ~ 1 pm.

The current design for LISA is to use heterodyning to develop accurate phase measurement. The required precision is $< 40\text{pm}/\sqrt{\text{Hz}}$. This precision is to be met by locking one laser to its reference cavity, while the other lasers are offset locked to it (two directly, three indirectly). The differing offset frequencies, set by oscillators on each spacecraft, will provide beat frequencies in the range of 75 to 125 kHz. Thus, LISA will be a virtual interferometer. It will not be measuring fringes as with a standard interferometer, but instead the phase differences will be combined on Earth. A Michelson signal can created by combining these phase differences:

$$S_1(t) = \Phi_{12}(t-L) + \Phi_{21}(t-L) - \Phi_{13}(t_3) - \Phi_{31}(t)\,, \tag{2.1}$$

where $\Phi_{ij}(t)$ is the phase difference between the signal transmitted by spacecraft i and received by spacecraft j at time t. However, Michelson-type combinations would be overcome by phase noise, so other combinations must be used. Some examples of more advanced combinations, such as the A and E channels [15], contain the detector response to gravitational waves, while another combination, known as the T-channel [15], is insensitive to low frequency gravitational waves, and can be used to monitor noise in the detector.

Noise Sources and Effects on Sensitivity

For interferometry the precise measurement of the phase of the laser light is necessary. The precision of these measurements is complicated by the presence of sources of experimental noise in the system. Noise in the case of LISA are effects and/or actions that cause fluctuations in these phase measurements. There are both internal and external sources of noise.

For space-based gravitational wave detectors, external sources of noise involve things such as solar wind, dust, debris, and possibly asteroids or comets passing near enough to the spacecraft to cause a non-negligible gravitational disturbance. Also, any of these items could pass through the path of the laser beam, possibly causing spurious signals.

Internal sources of noise involve uncertainties in the measurement apparatus. Laser (or position) noise is a general classification given to types of noise directly involving the laser system. Types of laser noise include photon shot noise and laser phase noise (due to power and phase fluctuations of the laser system), laser pointing noise (due to fluctuations in the directing of the laser beams), and clock noise (due to inaccuracies of the onboard timing system). Acceleration noise is the general classification of types of noise involving unexpected accelerations of the proof masses. Types of acceleration noise that are internal to the spacecraft include thermal noise (due to distortions of the spacecraft, the cavity containing the proof masses, and the telescope), electrical noise (due to electrical charging of spacecraft components by cosmic rays and Brownian motion due to dielectric losses), and magnetic noise (due to a proof mass, charged by cosmic rays, moving through the interplanetary magnetic field). These sources of noise can mimic changes in the optical path of the beam. Many expected sources will be too weak to be detectable above these noise sources, even with the best of data analysis techniques.

The size of the effect on the length of the optical path by significant sources of position noise are listed Table 2.1 [8]. These are caused by fluctuations in the laser and timing systems. Table 2.2 [8] lists the significant acceleration noise sources, also with the resulting error in measurement of the length of the optical path.

Table 2.1: Significant sources of noise in length measurement of the optical-path.

Error Source	Error Size ($\times 10^{-12}$ m/$\sqrt{\text{Hz}}$)	Number of Effects
Detector Shot Noise	11	4
Master Clock Noise	10	1
Residual Phase Noise	10	1
Laser beam-pointing instability	10	4
Laser phase measurement and offset lock	5	4
Scattered-light effects	5	4
Other substantial effects	3	32
Total path difference	40	= measurement error in $\mathcal{L}_2 - \mathcal{L}_1$

Table 2.2: Significant sources of acceleration noise.

Error Source	Error Size ($\times 10^{-15}$ m s$^{-2}/\sqrt{\text{Hz}}$)	Number of Effects
Thermal distortion of spacecraft	1	1
Thermal distortion of payload	0.5	1
Noise due to dielectric losses	1	1
Gravity noise due to spacecraft displacement	0.5	1
Temperature difference variations across the cavity	1	1
Electrical force on charged proof masses	1	1
Lorentz forces on charged proof masses	1	1
Residual gas impacts on proof masses	1	1
Thermal expansion of telescope	0.5	1
Magnetic forces on proof masses	0.5	1
Other substantial effects	0.5	4
Other smaller effects	0.3	16
Total effect of acceleration	3	for one inertial sensor
Effect in optical path	12	= variation in $\frac{\partial^2}{\partial t^2}\mathcal{L}_2 - \mathcal{L}_1$

Figure 2.4 shows the LISA sensitivity curve. The acceleration noise sources dominate at the lower end of the frequency range, but decrease as the square of the

frequency. Position noise, on the other hand, is relatively constant across the LISA band and surpasses the acceleration noise on the higher end of the frequency range. These effects appear below the LISA sensitivity curve due to averaging over all source polarizations and directions. Also affecting the higher end of the LISA sensitivity curve is the transfer function. The transfer function is due to interactions between the gravitational wave with the detector. Waves with frequencies above the transfer frequency ($f_* \equiv \frac{c}{2\pi L}$) will have periods less than the light crossing time of LISA's arms. This gives rise to a self-cancellation of the detected wave.

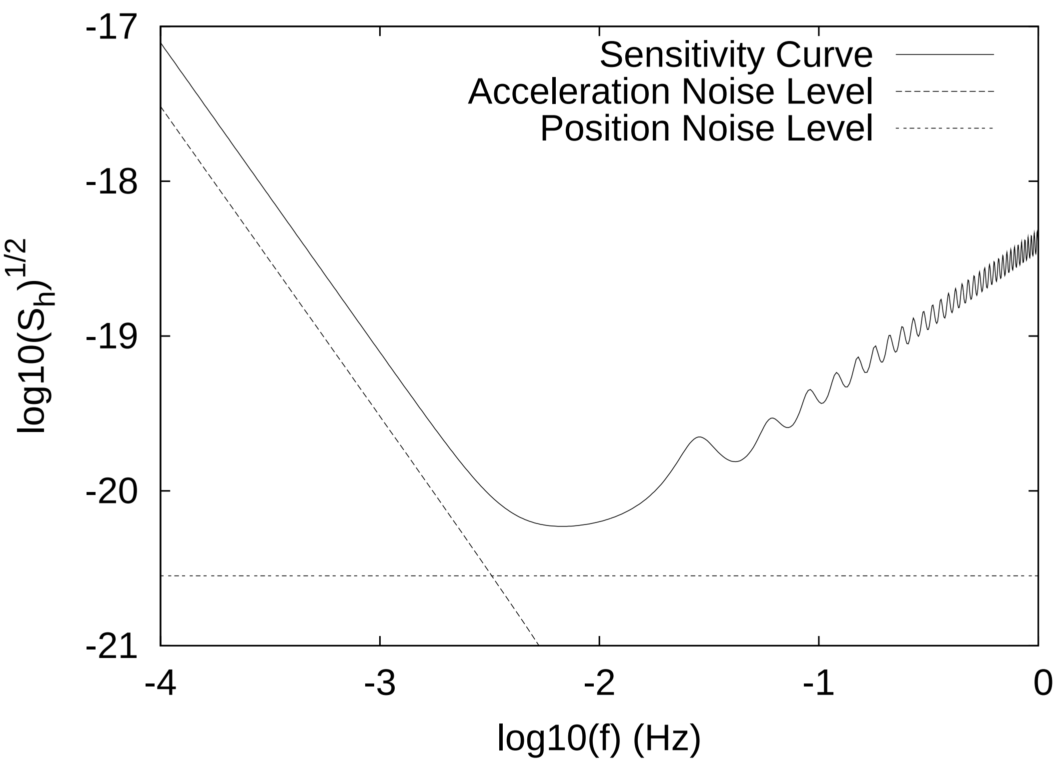

Figure 2.4: LISA sensitivity curve showing the instrumental sensitivity and the all-sky averaged overall sensitivity.

Modulations of the Signal

The orbital motion of LISA (shown in Figure 2.2) and the antenna pattern of LISA (shown in Figure 2.3) combine to modulate the gravitational waves passing through the detector. These modulations are called Doppler and amplitude modulations. They introduce variations into the signal due to a gravitational wave based on the location of the source.

Doppler modulation is caused when the orbital motion of LISA is moving the detector toward (or away from) a gravitational wave source. The movement of the detector increases (or decreases) the rate at which the wavefronts of the gravitational waves are passing through the antenna, creating a Doppler shift in the gravitational wave frequency. The amount of the shift at any given time depends on the relative position between detector and source. Thus, armed with the orbital specifics of LISA and a good model of the gravitational waveform produced by a source, one can arrive at the approximate location of the source.

Amplitude modulation is due to the shape of LISA's directional sensitivity combined with the $60°$ incline of the plane containing the three spacecraft with respect to the ecliptic. As LISA orbits over the course of a year, the sensitivity pattern varies across the sky. A gravitational wave source at any given location will undergo a modulation due to the variations in directional sensitivity.

The third modulation of the signal in a gravitational wave detector is known as phase modulation. This is different from Doppler modulation, which is a type of modulation of the gravitational wave phase. Phase modulation, in this case, is due to the presence of two polarizations in gravitational waves passing through the detector. The two polarizations have distinct amplitude modulations, which will introduce a varying phase difference between them.

These modulations affect all of the gravitational waves detected by LISA, but it is easiest to understand the effect by imagining a gravitational wave source continuously emitting at a specific frequency (such sources are labeled monochromatic). The Doppler, amplitude and phase modulation will spread the signal in the LISA data in multiples of its orbital frequency ($f_m = 1/\text{year}$), producing sidebands in the fourier transform of the signal. The bandwidths of the amplitude and phase modulations are $\sim 10^{-4}$ mHz. The bandwidths for Doppler modulation show a large range across the LISA band. However, for most sources, which will be emitting below 10 mHz, the bandwidths of equatorial sources will range from $\sim 8.3 \times 10^{-5}$ mHz (for $f_{source} = 0.1$ mHz) to 1.0×10^{-3} mHz (for $f_{source} = 10$ mHz).

Gravitational Wave Sources for LISA

This section will give a brief introduction to a number of general types of gravitational wave sources that are expected [16, 17, 18, 19] to be sources in the LISA band. These will provide an excellent means of testing General Relativity's predictions of the presence of gravitational waves, solidifying our knowledge of the astrophysics of the various system types, and providing detailed information about their spatial distributions. Yet, the sources that are known are only part of the picture. There will almost certainly also be unexpected gravitational wave sources. Part of the process in determining the nature of the unexpected sources will be the identification of the known sources, and the removal of their effects on the data, allowing a clearer look at how these unexpected sources affected LISA.

Galactic Binaries

While our galaxy is not particularly special in terms of the distribution of star systems contained inside it, those star systems do have the benefit of being close. Since the amplitudes of gravitational waves fall off as $1/r$, gravitational wave sources inside our galaxy will have much larger "apparent magnitudes." This will allow for LISA to detect many types of sources that will be undetectable in other galaxies. Examples of these "quiet" sources are neutron star (NS) binaries, close white dwarf binaries, stellar mass black hole binaries, helium cataclysmic variable binaries (made up of a helium star and a white dwarf), and "unevolved" binaries (made up of relatively young stars who are neither giants nor in the helium burning stage).

Though all the systems in each of these types of binaries is evolving toward a final coalescence through the loss of energy and angular momentum as gravitational waves are emitted, for many systems the pace of the evolution is so slow as to be negligible. These systems are the so-called monochromatic binaries, systems whose gravitational wave frequency is (effectively) unchanged during the time of observation.

A large segment of the galactic binaries that will be emitting gravitational waves in the low frequency regime will form what is sometimes referred to as the Gravitational Wave Confusion Background (GWCB). These binaries represent a possible barrier to the discovery of unexpected sources of gravitational waves. The GWCB and the nature of this barrier will be further discussed in Chapter 3.

Binary Mergers Involving Massive Black Holes

As was shown in Chapter 1, the magnitude of the gravitational waves generated by a binary depends directly on the product of the binary constituent masses. Thus binary systems where one or both objects is a massive black hole (MBH) will emit gravitational waves that will be detectible by LISA at great distances. For example, LISA would be able to detect a coalescing binary with masses $10^5 M_\odot$ and $1 M_\odot$ out to distances of ~ 1 Gpc, while a $10^5 M_\odot$ equal mass binary system merging anywhere in the Universe will be detectable.

As the two objects in the binary system spiral into each other, they increase their orbital and gravitational wave frequencies as well as the amplitude of the wave, providing a signature of the event known as the "chirp." The strongest emission of gravitational waves occurs at the end of the chirp, just before the actual merging of the binary constituents. This combination of signal strength and well modeled signature will allow for precise tests of General Relativity in the strong field regime. Detections will also provide information about the rate at which such events occur, and help test theories concerning black hole formation.

Gravitational Wave Background

The primordial Gravitational Wave Background (GWB) is an expected relic of the Big Bang, and is the least well understood of the proposed sources for LISA. A thermal GWB would have a temperature of ~ 0.9 K, and its amplitude in the LISA frequency band would be far below the LISA sensitivity curve. However, a non-thermal GWB may be detectable by LISA. Such a detection could provide information about early density fluctuations in the Universe. One possible cause that is receiving attention these days is the presence of cosmic superstrings [20, 21, 22].

Future Space-Based Gravitational Wave Detectors

Though LISA is still about a decade from launch, space-based follow-on missions are already being proposed. They each share the same basic operational design of LISA, interfometers involving three spacecraft or multiple groupings of three spacecraft, orbiting the Sun. Capabilities of the the follow-on missions detailed here are examined in detail in Chapter 4.

Laser Interferometer Space Antenna in Stereo

The Laser Interferometer Space Antenna in Stereo (LISAS) [23] is the simplest follow-on to LISA. While LISA consists of one constellation of three spacecraft in a 20° Earth-trailing orbit, LISAS would consist of two such constellations, one in the 20° Earth-trailing orbit and another in a 20° Earth-leading orbit. The parameters for the components and configurations of each of the spacecraft (and constellations of spacecraft) would be identical to LISA. Thus should the LISA mission outlive its current five year lifetime, the LISAS mission could be implemented by the launch of just the second constellation. The benefit gained by the presence of the second, separated constellation is added angular resolution, which is greatly enhanced for coalescing MBH binaries. The exact nature of these improvements is shown in Chapter 4

Advanced Laser Interferometer Antenna and Advanced Laser Interferometer Antenna in Stereo

The Advanced Laser Interferometer Antenna (ALIA) [24] is a single constellation of three spacecraft as in the LISA configuration, with the spacecraft separation only one-tenth that of LISA (5×10^8 m). Also, the proposed level for the acceleration noise of ALIA is one-tenth that of LISA, while the proposed position noise level is one two-hundredth of LISA. These changes increase the sensitivity at higher frequencies compared to that of LISA (or LISAS). This allows ALIA more detailed observations involving coalescing intermediate mass black holes (IMBHs) with masses in the range of $50M_\odot$ to $50,000M_\odot$. This information will help test theories of IMBH formation and population distribution.

The Advanced Laser Interferometer Antenna in Stereo (ALIAS) [23] is two constellations of three spacecraft with the same parameters as ALIA. The two constellations are separated by 40° (one trailing and one leading the Earth by 20°). As with LISAS the second, separated constellation provides a much improved resolution for the coalescence of IMBHs (see Chapter 4 for details) beyond the capabilities of ALIA.

Big Bang Observer

The Big Bang Observer (BBO) [25] is another follow-on mission with up to four constellations of three spacecraft each, with armlengths one-hundredth that of LISA

(5×10^7 m). Similarly its acceleration noise budget is one-hundredth that of LISA, and its position noise budget is one-millionth that of LISA. The BBO is designed to detect GWB created shortly after the Big Bang, during the inflationary period of the Universe. The orbital configuration of the constellations is shown in Figure 2.5.

The initial phase of the BBO is two constellations in a 20° Earth-trailing orbit. The two constellations overlap each other forming the shape of a six pointed star. This design allows for optimal cross-correlation of the gravitational wave signals between the constellations, with a minimum of correlated noise [14]. Thus, over the observation time, the correlated signals will grow while the uncorrelated noise will tend to average toward zero. This will allow for the extremely sensitive levels of detection that will be necessary to measure the GWB.

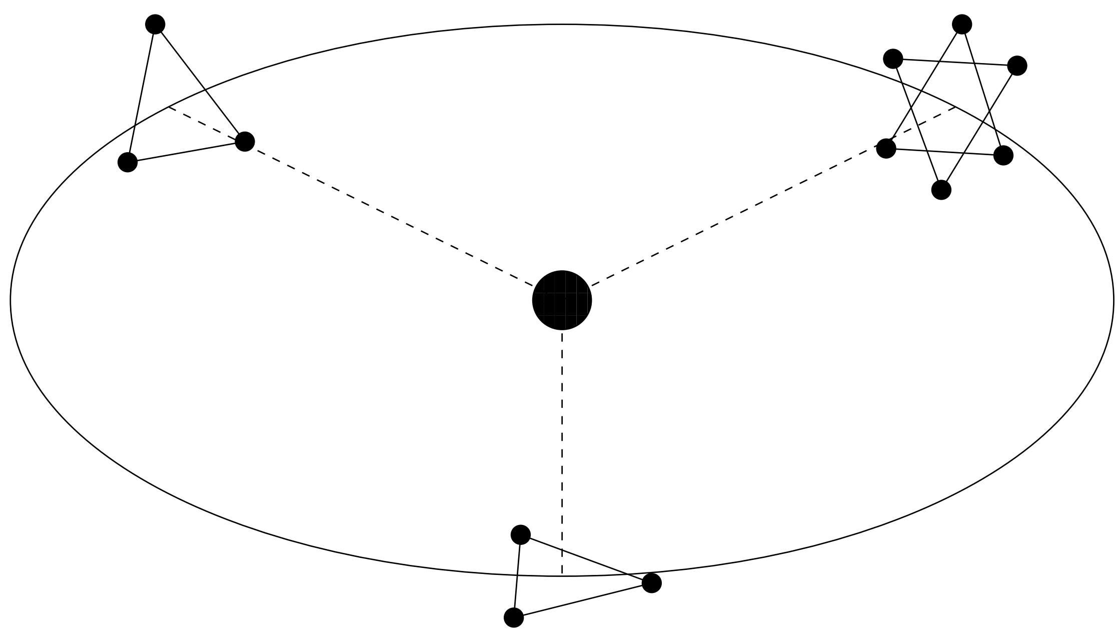

Figure 2.5: The proposed orbital configuration of the Big Bang Observer.

The next phase of the BBO involves the addition of two more constellations, separated from the star and each other by 120° in the plane of the ecliptic around the Sun. These constellations provide remarkable resolution capabilities for coalescing binary systems. The initial expectation was that the extreme resolution would be needed in order to accurately remove the signals due to low amplitude sources such as NS-NS binary systems. However, recent studies [26] bring the need for such extremes in resolution into question.

CHAPTER 3

DATA ANALYSIS FOR GRAVITATIONAL WAVES

Introduction

In scientific experiments, generally speaking, there are two basic questions to be answered. The first question is, 'what is the measured value of the result being sought.' The second question is, 'to what precision is that the result known.' The field of data analysis seeks to provide the answers to these questions, which for LISA means working backward from the data to the parameters of the gravitational wave sources. Due to the presence of noise in the data this "inversion problem" is ill-posed, and data analysis techniques will be required to obtain those parameters. While the field of data analysis is quite broad and includes numerous techniques and methods, the scope of this chapter will be limited to those techniques and methods being used to prepare for answering those two basic questions concerning the data from space-based gravitational wave detectors.

The heart of LISA data is a series of phase measurements of the laser light that will be used to derive the corresponding strains of the detectors arms. By modeling the gravitational waveforms produced by a source the strain of the detector arms can be predicted. Matching prediction to the data will allow for detection.

There are various approaches to statistical inference in data analysis. One such approach is the frequentist approach. In terms of gravitational wave detection this is exemplified by repeated measurements of a well-defined experiment involving random noise. One type of a frequentist's approach to measuring detector resolution involves repeated attempts at extracting the parameters of a specific binary system under multiple noise realizations. While this can help in understanding the resolution of a detector before it is operational, the measurements that functioning space-based detectors will be making are of an entirely different nature. There will be no chance to sample different noise realizations.

Another approach is the Bayesian approach. Bayes' theorem, which lays at the heart of this approach, tells how one is to update or change the degree of belief held in a proposition with the introduction of new evidence. The updating process is known as Bayesian inference, and it is a statistical analog to the scientific method. Thus, this approach will closely model the operational period of gravitational wave detectors.

What will be used in this work is labeled by some as the "eclectic" approach, in that both frequentist and Bayesian approaches will be used in the cases they are best fit to handle. Chapters 4 and 5 use a frequentist approach, determining the resolving power of various detectors under differing conditions. Chapters 6 and 7 use a Bayesian approach, searching specific data streams to extract the parameters of the gravitational wave sources used to create the data.

The succeeding chapters, which will use the techniques described below, follow the progression of data analysis needs for the development of a detector. First, one needs to understand the detector's capabilities. In Chapter 4 an investigation of detector resolution for a single source is presented. When required precision in a simple case is assured, complexity is added to the model. Chapter 5 provides a study of the effects on parameter resolution due to the presence of multiple gravitational wave source. Next, one must know if the parameters being sought will be able to be determined form the detector's data. Chapter 6 contains a survey of a search method that will extract source parameter values from a data snippet with several sources. And Chapter 7 develops an algorithm that will both search for and characterize the distributions of the source parameter from a data snippet containing multiple binary systems.

Before delving into specific techniques, a few useful definitions are provided. The noise weighted inner product is given by:

$$(a|b) = 2\int_0^\infty \frac{\tilde{a}^*(f)\tilde{b}(f) + \tilde{a}(f)\tilde{b}^*(f)}{S_n(f)}\, df, \tag{3.1}$$

where a and b are gravitational waveforms, an asterisk denotes the waveform's complex conjugate, a tilde denotes its Fourier transform, and $S_n(f)$ is the noise spectral density. Thus the signal to noise ratio (SNR) is defined as:

$$\mathrm{SNR}^2 = \sum_\alpha (h_\alpha|h_\alpha)\,, \tag{3.2}$$

where the inner product of the gravitational waves, h_α, is summed over the independent channels of the detector.

Also, one can write the output the LISA data stream from detector α as:

$$s_\alpha(t) = h_\alpha(t,\vec{\lambda}) + n_\alpha(t) = \sum_{i=1}^{N} h_\alpha^i(t,\vec{\lambda}_i) + n_\alpha(t)\,, \tag{3.3}$$

where $\vec{\lambda}$ is the complete set of parameters describing the N sources of the superposed gravitational waves, $h_\alpha^i(t,\vec{\lambda}_i)$, whose individual parameters are $\vec{\lambda}_i$, and $n_\alpha(t)$ is the detector noise.

Matched Filtering

The inherent weakness of the interaction between the spacecraft (or any mass) and gravitational waves will mandate a method to pull many of the weaker signals out from the detector noise. The optimal method, called matched filtering, requires a known model for the gravitational wave source. Matched filtering is the primary data analysis technique of modeled sources for LISA, and may be used in conjunction with other techniques. Matched filtering is used extensively in all fields of science, and is a popular data analysis technique in ground based gravitational wave astronomy [27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38].

The LISA data can be filtered through the various models, and source parameter values, to extract the matches that may exist. By filtering, or projecting, the signal onto a specific model and parameter set, called a template, one can get a measure of their agreement or overlap. If the noise in the detector is gaussian with a zero mean, the probability, also known as the likelihood, of such a match between a signal, s, and a gravitational wave from a source with parameters, $\vec{\lambda}$, is given by [39]:

$$p(s|\vec{\lambda}) = C \exp\Big[-\frac{1}{2}\big((s - h(\vec{\lambda}))|(s - h(\vec{\lambda}))\big)\Big] \tag{3.4}$$

where C is a normalization constant, independent of both signal and template.

One flavor of best fit parameters is found by maximizing the likelihood (or alternatively the SNR). Another measure whose maximization gives the best fit parameters, will be referred to as the log likelihood (though it is truly involves an inversion and rescaling of the logarithm of (3.4)), is given by:

$$\log \mathcal{L} = (s|h(\vec{\lambda})) - \frac{1}{2}(h(\vec{\lambda})|h(\vec{\lambda})). \tag{3.5}$$

Note that if the gravitational wave signals is uncorrelated with the noise:

$$\langle \log \mathcal{L}(\vec{\lambda}) \rangle = \frac{1}{2}\,\mathrm{SNR}^2\,. \tag{3.6}$$

In this section the angle brackets $\langle\rangle$ denote an expectation value.

To search for possible sources hidden in the noise, one could imagine matching distinct templates that have been distributed across the space of parameters. Such grid-based template searches are reliable, but for LISA, where parameters for single sources will number as high as 17, are expected to be computationally prohibitive without marked improvements in either computer speed or efficiency of template spacing.

Fisher Information Matrix

While a template search using matched filtering is a global technique, allowing for the measurement of matches across the parameter space, the Fisher Information Matrix (FIM) approach is specific to a localized area around the best fit parameter values. It is one example of a Maximum Likelihood (ML) Estimation Method, in which the ML estimator maximizes ratio of the likelihood (3.4) to the prior probability distribution of the data set, $p(\vec{x})$. That is the ML estimator is the parameter set that simultaneously solves the ML equation,

$$\frac{\partial\Lambda(\vec{\lambda},\vec{x})}{\partial\lambda_i} = 0\,, \tag{3.7}$$

for each of the parameters, i, in the model, with $\Lambda(\vec{\lambda},\vec{x}) \equiv \frac{p(s(\vec{x})|\lambda)}{p(\vec{x})}$. Please note that there is not universal agreement on the definition of the ML estimator. The definition used here is the standard definition used by statisticians and some [40, 41] in the gravitational wave community. Others [31, 39] refer to the dominant mode of the posterior distribution, $p(\vec{\lambda}|s)$, as a maximum likelihood estimator. The latter has the advantage of incorporating prior information, but the disadvantage of not being invariant under parameter space coordinate transformations. The former definition, while it does not take into account prior information, is coordinate invariant. The two definitions give the same result for uniform priors, and very similar results in most cases (the exception being where the priors have a large gradient at maximum likelihood).

The FIM, Γ, is given as the negative of the expectation value of the Hessian evaluated at maximum likelihood:

$$\Gamma_{ij} = -\left\langle \frac{\partial^2 \log \mathcal{L}}{\partial \lambda^i \partial \lambda^j} \right\rangle_{\rm ML} = (h_{,i}|h_{,j})_{\rm ML}\,, \tag{3.8}$$

where $h_{,i} \equiv \frac{\partial h}{\partial \lambda^i}$. The variance-covariance matrix, C is estimated by the inverse of Γ. The lower bound on the variances are given by the Cramer-Rao bound.

$$C_{ii} \geq (\Gamma^{-1})_{ii}. \tag{3.9}$$

The first order error in this estimate is inversely proportional to the SNR of the source. For large SNR the estimated parameter uncertainties, $\Delta\lambda^i$, will have the Gaussian probability distribution given by [31]:

$$p(\Delta\lambda^i) = \sqrt{\frac{\det\Gamma}{2\pi}} \exp\left(-\frac{1}{2}\Gamma_{ij}\Delta\lambda^i\Delta\lambda^j\right)\,. \tag{3.10}$$

As the variance-covariance,

$$C_{ij} = \langle \Delta\lambda^i \Delta\lambda^j \rangle\,. \tag{3.11}$$

provides estimates for the parameter uncertainties and correlation, the square root of the diagonal elements of the variance-covariance matrix, $\Delta\lambda^i = (C^{ii})^{1/2}$, will be used to evaluate parameter uncertainties with the FIM approach. From equations (3.5) and (3.8) it follows that the parameter uncertainties scale inversely with the SNR.

Parameter Estimation

The discussion here is applied to a single monochromatic binary system. Chapter 4 applies the method to coalescing binaries (which requires only a change of the parameter set which grows from seven to nine parameters), and Chapter 5 extends the method to multiple sources.

Typical galactic binaries can be treated as circular and monochromatic. They are described by seven parameters: sky location (θ, ϕ); gravitational wave frequency f; amplitude A; inclination and polarization angles (ι, ψ); and the initial orbital phase φ_0. The response of a space-borne detector to a gravitational wave source is encoded in the (Micheleson-like) $X_i(t)$ time-delay interferometry variables [42]. Here the subscript i denotes the vertex at which the signal is read out. This work employs the rigid adiabatic approximation [43] to describe these variables.

If the analysis to be done is limited to low frequencies, $f \ll f_* \equiv \frac{c}{2\pi L}$, where L is the length of the detector arms, one can work with the orthogonal combinations such as [44]:

$$S_I(t) = X_1(t), \quad S_{II}(t) = \frac{1}{\sqrt{3}} \left(X_1(t) + 2X_2(t) \right) . \tag{3.12}$$

If higher frequencies are involved, in the equal-armlength limit, the TDI signal $X_i(t)$ can be formed from a time-delayed differencing of Michelson signals $M_i(t)$ that will cancel the laser phase noise, while preserving the gravitational wave signal:

$$X_i(t) = M_i(t) - M_i(t - 2L) . \tag{3.13}$$

Thus one can use orthogonal combinations such as [45]:

$$\begin{aligned} A(t) &= \sqrt{\frac{3}{2}} \left(S_a(t) - S_a(t - L) \right) , \\ E(t) &= \sqrt{\frac{3}{2}} \left(S_e(t) - S_e(t - L) \right) , \end{aligned} \tag{3.14}$$

where

$$S_e(t) = \frac{1}{3}\left(2M_1(t) - M_2(t) - M_3(t)\right),$$
$$S_a(t) = \frac{1}{\sqrt{3}}\left(M_2(t) - M_3(t)\right). \tag{3.15}$$

The A, E combinations cancel the laser phase noise that would otherwise dominate the Michelson signals. At high frequencies, $f > f_*$, a third independent data channel $T(t)$ becomes available. For frequencies below f_* the T-channel is insensitive to gravitational waves and can be used as an instrument noise monitor. To simplify the analysis this work does not use the T-channel. Including this channel would not have any great effect on the results as most of the signal to noise ratio accumulates below the respective transfer frequencies of the detectors considered. In the low frequency limit the variables $A(t)$ and $E(t)$ reduce to the independent data channels defined by (3.12).

In the low frequency limit the signal of a monochromatic source can be written [44]:

$$s_1(t) = A(t)\cos[\Phi(t; f, \theta, \phi) + \varphi_0]\,, \tag{3.16}$$

where

$$A(t) = [A_+^2 F^{+2}(t) + A_\times^2 F^{\times 2}(t)]^{1/2} \tag{3.17}$$

is the amplitude of the signal and the gravitational wave phase is

$$\Phi(t; f, \theta, \phi) = 2\pi f t + \phi_D(t; f, \theta, \phi) + \phi_P(t, \theta, \phi), \tag{3.18}$$

While

$$\phi_D(t) = 2\pi f \mathrm{AU} \sin\theta \cos\left(2\pi f_m t - \phi\right) \tag{3.19}$$

is the phase shift due to Doppler modulation and $f_m = 1/\text{year}$ is the modulation frequency, and

$$\phi_P(t) = -\arctan\Big(\frac{A_\times F^\times(t)}{A_+ F^+(t)}\Big) \tag{3.20}$$

is the phase modulation due to combining the two polarizations of the gravitational wave.

The beam pattern factors, $F^+(t)$ and $F^\times(t)$, in these equations are given by:

$$\begin{aligned} F^+(t) &= \frac{1}{2}\left(\cos 2\psi\, D^+(t) - \sin 2\psi\, D^\times(t)\right) \\ F^\times(t) &= \frac{1}{2}\left(\sin 2\psi\, D^+(t) + \cos 2\psi\, D^\times(t)\right)\,, \end{aligned} \tag{3.21}$$

where ψ is the polarization angle of the gravitational wave source, the amplitudes of the two polarizations are:

$$\begin{aligned} A_+ &= (1+\cos^2\iota)A \\ A_\times &= -2\cos\iota A\,, \end{aligned} \tag{3.22}$$

and the detector pattern functions $D^+(t)$ and $D^\times(t)$ are to linear order [43] in the eccentricity of Keplerian orbits:

$$\begin{aligned} D^+(t) = \frac{\sqrt{3}}{64}\Bigg[&-36\sin^2(\theta)\sin\big(2\alpha(t)-2\lambda\big) + \big(3+\cos(2\theta)\big)\Bigg(\cos(2\phi)\Big(9\sin(2\lambda) \\ &-\sin\big(4\alpha(t)-2\lambda\big)\Big) + \sin(2\phi)\Big(\cos\big(4\alpha(t)-2\lambda\big) - 9\cos(2\lambda)\Big)\Bigg) \\ &-4\sqrt{3}\sin(2\theta)\Big(\sin\big(3\alpha(t)-2\lambda-\phi\big) - 3\sin\big(\alpha(t)-2\lambda+\phi\big)\Big)\Bigg] \end{aligned} \tag{3.23}$$

$$\begin{aligned} D^\times(t) = \frac{1}{16}\Bigg[&\sqrt{3}\cos(\theta)\Big(9\cos\big(2\lambda-2\phi\big) - \cos\big(4\alpha(t)-2\lambda-2\phi\big)\Big) \\ &-6\sin(\theta)\Big(\cos\big(3\alpha(t)-2\lambda-\phi\big) + 3\cos\big(\alpha(t)-2\lambda+\phi\big)\Big)\Bigg]\,. \end{aligned} \tag{3.24}$$

Applying this formalism to an isolated galactic binary yields the parameter uncertainties shown in Figures 3.1 and 3.2. The plots were generated by taking the median values for 10^5 different sources, each normalized to a signal-to-noise ratio of SNR = 10. The uncertainty in the frequency has been increased by a factor of 10^6 for plotting purposes.

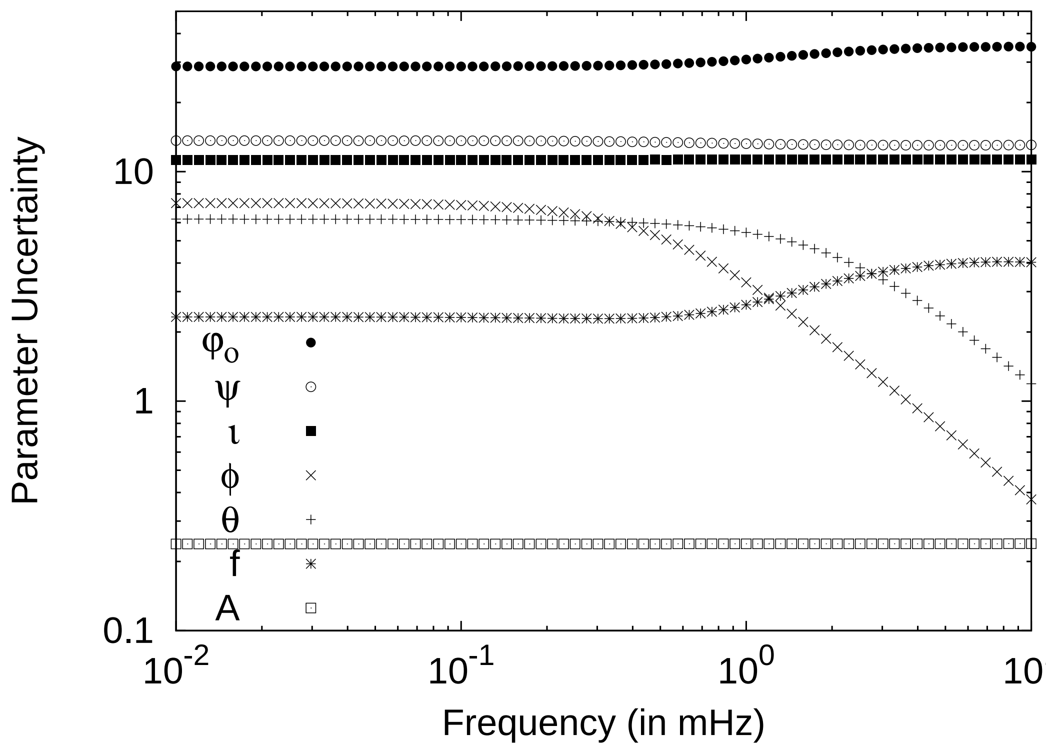

Figure 3.1: Median parameter uncertainties versus frequency for isolated monochromatic binary sources with one year of observation and fixed signal-to-noise ratios of SNR = 10. The parameters are $f, \ln A, \theta, \phi, \iota, \psi, \varphi_0$.

As can be seen in Figure 3.1, the uncertainties in five of the seven parameters are fairly constant across the LISA band. The two exceptions are the sky location variables (θ, ϕ), which show a marked decrease in their uncertainties above 1 mHz. This behavior can be traced to the time varying Doppler shift, which is one of the two ways LISA is able to locate sources on the sky. The Doppler shift increases linearly with f, which translates into a $1/f$ decrease in the positional uncertainties above

1 mHz. Below this frequency the angular resolution comes mainly from the time varying antenna sweep, which is weakly dependent on f below the transfer frequency $f_* \simeq 10$ mHz.

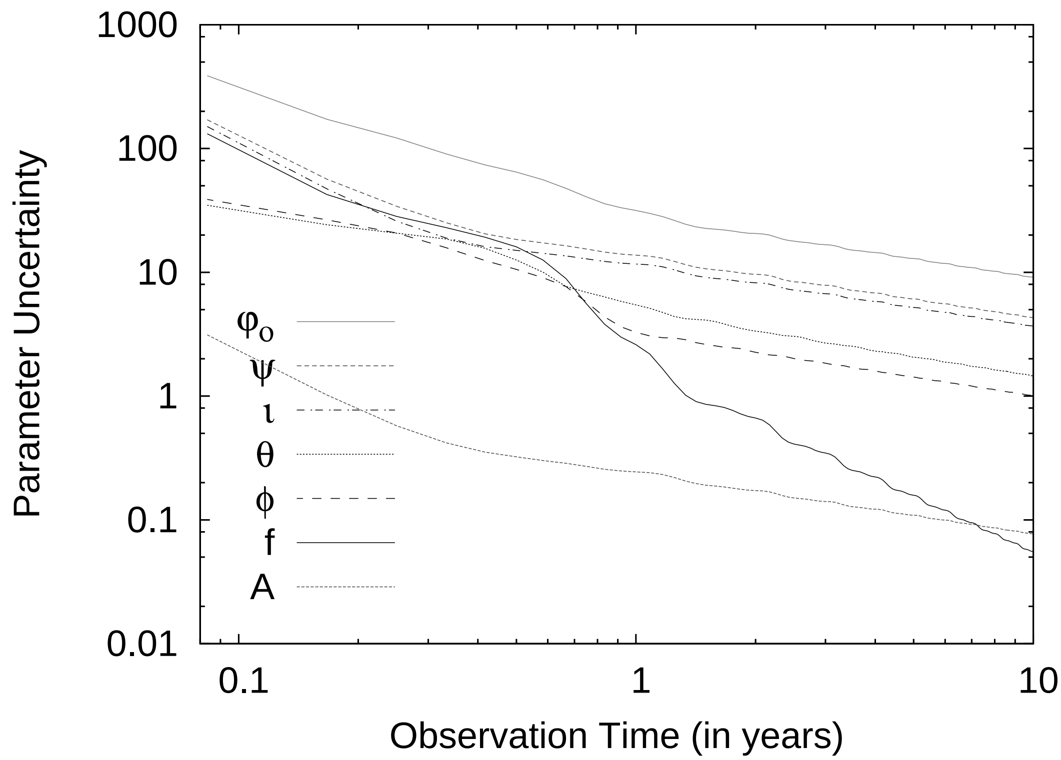

Figure 3.2: Parameter uncertainties versus time for isolated monochromatic binary sources with $f = 1$ mHz and fixed signal-to-noise ratios of SNR $= 10$. The parameters are $\ln f, \ln A, \theta, \phi, \iota, \psi, \varphi_0$.

As can be seen in Figure 3.2, the uncertainties in six of the seven parameters decrease as $1/\sqrt{T_{obs}}$ for observation times longer than a year, in accordance with the standard $\sqrt{T_{obs}}$ increase in the SNR. The one exception is the frequency, which benefits from an additional $1/T_{obs}$ shrinkage in the width of the frequency bins, leading to an overall $1/T_{obs}^{3/2}$ decay in the frequency uncertainty. The shrinkage in the size of the frequency bins is illustrated in Figure 3.3, where the discrete Fourier transform of a typical signal is shown after one year and after ten years of observation. In Chapter 5 it will be seen how this improved resolution of the sidebands helps reduce source confusion when multiple overlapping sources are present. For observation times less than one year the positional uncertainties also decay as $1/T_{obs}^{3/2}$. The additional factor of $1/T_{obs}$ can be traced to the two mechanisms (modulations) by which LISA can locate a source: antenna sweep and Doppler shift. The angular resolution due to antenna sweep depends on the angle through which the antenna has been swept, which increases as T_{obs} for $T_{obs} <$ year. The angular resolution due to the variable Doppler shift can be thought of as a form of long baseline interferometery. The resolution improves with the length of the synthesized baseline, which grows as T_{obs} for $T_{obs} <$ year.

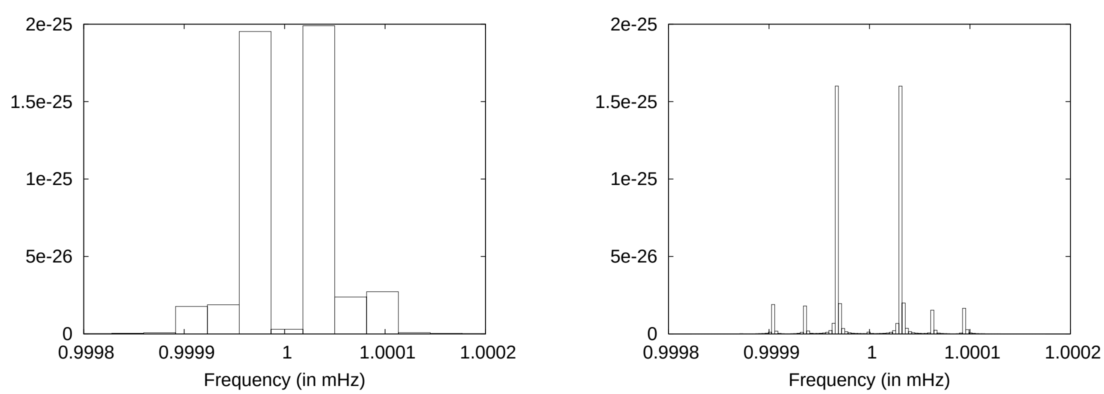

Figure 3.3: Fourier transforms of a typical signal for observation times of one and ten years.

The Cocktail Party Problem

When space-based detectors are operational they will be affected by millions of low SNR sources at frequencies below 3 mHz.

Extracting the parameters of each individual source from the combined response to all sources defines the LISA cocktail party problem. In practice it will be impossible to resolve all of the millions of signals that contribute to the LISA data streams. For one, there will not be enough bits of information in the entire LISA data archive to describe all N sources in the Universe with signals that fall within the LISA band. Moreover, most sources will produce signals that are well below the instrument noise level, and even after optimal filtering most of these sources will have signal to noise ratios below one. A more reasonable goal might be to provide estimates for the parameters describing each of the N' sources that have integrated signal to noise ratios (SNR) above some threshold (such as SNR > 5), where it is now understood that the noise includes the instrument noise, residuals from the regression of bright sources, and the signals from unresolved sources. While the noise will be neither stationary nor Gaussian, it is not unreasonable to hope that the departures from Gaussianity and stationarity will be mild. It is well know that matched filtering is the optimal linear signal processing technique for signals with stationary Gaussian noise [46, 47].

In the Fourier domain, the signal can be written as $\tilde{s}_\alpha(f) = \tilde{h}_\alpha(f, \vec{\lambda}') + \tilde{n}_\alpha(f)$, where $\tilde{n}_\alpha(f)$ includes instrument noise and confusion noise, and the signals are described by parameters $\vec{\lambda}'$. Using the standard noise weighted inner product for the independent data channels over a finite observation time T,

$$(a|b) = \frac{2}{T}\sum_\alpha \sum_f \frac{\tilde{a}^*_\alpha(f)\tilde{b}_\alpha(f) + \tilde{a}_\alpha(f)\tilde{b}^*_\alpha(f)}{S^\alpha_n(f)}\,, \tag{3.25}$$

a Wiener filter statistic can be defined:

$$\rho(\vec{\lambda}) = \frac{(s|h(\vec{\lambda}))}{\sqrt{(h(\vec{\lambda})|h(\vec{\lambda}))}}\,. \tag{3.26}$$

The noise spectral density $S_n(f)$ is given in terms of the autocorrelation of the noise

$$\langle n(f) n^*(f') \rangle = \frac{T}{2}\delta_{ff'} S_n(f)\,. \tag{3.27}$$

The number of parameters d_i required to describe a source ranges from 7 for a slowly evolving circular galactic binary to 17 for a massive black hole binary. A reasonable estimate [48] for N' is around 10^4, so the full parameter space has dimension $D = \sum_i d_i \sim 10^5$. Since the number of templates required to uniformly cover a parameter space grows exponentially with D, a grid based search using the full optimal matched filter is out of the question. Clearly an alternative approach has to be found. Moreover, the number of resolvable sources N' is not known a priori, so some stopping criteria must be found to avoid over-fitting the data.

Most approaches to the LISA cocktail party problem employ iterative schemes. The first such approach was dubbed “gCLEAN” [49] due to its similarity with the “CLEAN” [50] algorithm that is used for astronomical image reconstruction. The “gCLEAN” procedure identifies and records the brightest source that remains in the data stream, then subtracts a small amount of this source. The procedure is iterated until a prescribed residual is reached, at which time the individual sources are reconstructed from the subtraction record. A much faster iterative approach dubbed “Slice & Dice” [51] was recently proposed that proceeds by identifying and fully subtracting the brightest source that remains in the data stream. A global least squares re-fit to all the current list of sources is then performed, and the new parameter record is used to produce a regressed data stream for the next iteration. Bayes factors are used to provide a stopping criteria.

There is always the danger with iterative approaches that the procedure “gets off on the wrong foot,” and is unable to find its way back to the optimal solution. This can happen when two signals have a high degree of overlap. A very different approach to the LISA source confusion problem is to solve for all sources simultaneously using ergodic sampling techniques, such as the Markov Chain Monte Carlo (MCMC) [52, 53, 54] method or Genetic Algorithms (GAs) [55, 56]. While both of these approaches provide a means for searching the data for optimal parameter fits with very large parameter spaces, MCMC methods also allow for an estimate of the posterior distribution, $p(\vec{\lambda}|s)$.

The MCMC method is now in widespread use in many fields, and is starting to be used by astronomers and cosmologists. One of the advantages of MCMC is that it combines detection, parameter estimation, and the calculation of confidence intervals in one procedure, as everything one can ask about a model is contained in the posterior distribution. Another nice feature of MCMC is that there are implementations that allow the number of parameters in the model to be variable, with built-in penalties for using too many parameters in the fit. In an MCMC approach, parameter estimates from Wiener matched filtering are replaced by the Bayes estimator [57]

$$\lambda^i_{\rm B}(s) = \int \lambda^i \, p(\vec{\lambda}|s) \, d\vec{\lambda} \,, \tag{3.28}$$

which requires knowledge of $p(\vec{\lambda}|s)$ - the posterior distribution of $\vec{\lambda}$ (*i.e.* the distribution of $\vec{\lambda}$ conditioned on the data s). By Bayes theorem, the posterior distribution is related to the prior distribution $p(\vec{\lambda})$ and the likelihood $p(s|\vec{\lambda})$ by

$$p(\vec{\lambda}|s) = \frac{p(\vec{\lambda})p(s|\vec{\lambda})}{\int p(\vec{\lambda}')p(s|\vec{\lambda}') \, d\vec{\lambda}'} \,. \tag{3.29}$$

Until recently the Bayes estimator was little used in practical applications as the integrals appearing in (3.28) and (3.29) are often analytically intractable. The traditional solution has been to use approximations to the Bayes estimator, such as the maximum likelihood estimator described above, however advances in the MCMC technique allow direct numerical estimates to be made. Recalling that when the noise $n(t)$ is a normal process with zero mean, the likelihood is given by (3.4), and in the large SNR limit the Bayes estimator can be approximated by finding the dominant mode of the posterior distribution, $p(\vec{\lambda}|s)$.

The expectation value of the maximum of the log likelihood is:

$$\langle \log \mathcal{L}(\vec{\lambda}_{\rm ML}) \rangle = \frac{{\rm SNR}^2 + D}{2} \, . \tag{3.30}$$

This value exceeds that found in (3.6) by an amount that depends on the total number of parameters used in the fit, D, reflecting the fact that models with more parameters generally give better fits to the data. Deciding how many parameters to allow in the fit is an important issue in LISA data analysis as the number of resolvable sources is not known a priori. This issue does not usually arise for ground based gravitational wave detectors as most high frequency gravitational wave sources are transient. The relevant question there is whether or not a gravitational wave signal is present in a section of the data stream, and this question can be dealt with by the Neyman-Pearson test or other similar tests that use thresholds on the likelihood, $\mathcal{L}$, that are related to the false alarm and false dismissal rates. Demanding that $\mathcal{L} > 1$ - so it is more likely that a signal is present than not - and setting a detection threshold of $\rho = 5$ yields a false alarm probability of 0.006 and a detection probability of 0.994 (if the noise is stationary and Gaussian). A simple acceptance threshold of $\rho = 5$ for each individual signal used to fit the LISA data would help restrict the total number of parameters in the fit, however there are better criteria that can be employed. The simplest is related to the Neyman-Pearson test and compares the likelihoods of models with different numbers of parameters. For nested models this ratio has an approximately chi squared distribution which allows the significance of adding extra parameters to be determined from standard statistical tables. A better approach is to compute the Bayes factor,

$$B_{XY} = \frac{p_X(s)}{p_Y(s)} \, , \tag{3.31}$$

which gives the relative weight of evidence for models X and Y in terms of the ratio of marginal likelihoods

$$p_X(s) = \int p(s|\vec{\lambda}, X)p(\vec{\lambda}, X)\, d\vec{\lambda}\,. \tag{3.32}$$

Here $p(s|\vec{\lambda}, X)$ is the likelihood distribution for model X and $p(\vec{\lambda}, X)$ is the prior distribution for model X. The difficulty with this approach is that the integral in (3.32) is hard to calculate, though estimates can be made using the Laplace approximation or the Bayesian Information Criterion (BIC) [58]. The Laplace approximation is based on the method of steepest descents, and for uniform priors yields

$$p_X(s) \simeq p(s|\vec{\lambda}_{\rm ML}, X)\left(\frac{\Delta V_X}{V_X}\right)\,, \tag{3.33}$$

where $p(s|\vec{\lambda}_{\rm ML}, X)$ is the maximum likelihood for the model, V_X is the volume of the model's parameter space, and ΔV_X is the volume of the uncertainty ellipsoid (estimated using the FIM). Models with more parameters generally provide a better fit to the data and a higher maximum likelihood, but they get penalized by the $\Delta V_X/V_X$ term which acts as a built-in Occam's razor.

Generalized F-Statistic

The F-statistic was originally introduced [41] in the context of ground based searches for gravitational wave signals from rotating NSs. The F-statistic has since been used to search for monochromatic galactic binaries using simulated LISA data [51, 59]. By using multiple linear filters, the F-statistic is able to automatically extremize the log likelihood over extrinsic parameters, thus reducing the dimension of the search space (the parameter space dimension remains the same).

In the low-frequency limit the LISA response to a gravitational wave with polarization content $h_+(t)$, $h_\times(t)$ can be written as

$$h(t) = h_+(t)F^+(t) + h_\times(t)F^\times(t)\,. \tag{3.34}$$

$$h(t) = \sum_{i=1}^{4} a_i(A, \psi, \iota, \varphi_0)A^i(t; f, \theta, \phi)\,, \tag{3.35}$$

where the time-independent amplitudes a_i are given by

$$\begin{aligned}
a_1 &= \frac{A}{2}\left((1+\cos^2\iota)\cos\varphi_0\cos 2\psi - 2\cos\iota\sin\varphi_0\sin 2\psi\right), \\
a_2 &= -\frac{A}{2}\left(2\cos\iota\sin\varphi_0\cos 2\psi + (1+\cos^2\iota)\cos\varphi_0\sin 2\psi\right), \\
a_3 &= -\frac{A}{2}\left(2\cos\iota\cos\varphi_0\sin 2\psi + (1+\cos^2\iota)\sin\varphi_0\cos 2\psi\right), \\
a_4 &= \frac{A}{2}\left((1+\cos^2\iota)\sin\varphi_0\sin 2\psi - 2\cos\iota\cos\varphi_0\cos 2\psi\right),
\end{aligned} \tag{3.36}$$

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

$$\begin{aligned}
A^1(t) &= D^+(t;\theta,\phi)\cos\Phi(t;f,\theta,\phi) \\
A^2(t) &= D^\times(t;\theta,\phi)\cos\Phi(t;f,\theta,\phi) \\
A^3(t) &= D^+(t;\theta,\phi)\sin\Phi(t;f,\theta,\phi) \\
A^4(t) &= D^\times(t;\theta,\phi)\sin\Phi(t;f,\theta,\phi)\,.
\end{aligned} \tag{3.37}$$

Defining the four constants $N^i = (s|A^i)$ and using (3.35) yields a solution for the amplitudes a_i:

$$a_i = (M^{-1})_{ij}N^j\,, \tag{3.38}$$

where $M^{ij} = (A^i|A^j)$. The output of the four linear filters, N^i, and the 4×4 matrix M^{ij} can be calculated using the same fast Fourier space techniques [48] used to generate the full waveforms. Substituting (3.35) and (3.38) into expression (3.5) for the log likelihood yields the F-statistic

$$\mathcal{F} = \log \mathcal{L} = \frac{1}{2}(M^{-1})_{ij} N^i N^j \,. \tag{3.39}$$

The F-statistic automatically maximizes the log likelihood over the extrinsic parameters A, ι, ψ and φ_0, and reduces the search to the sub-space spanned by f, θ and ϕ. The extrinsic parameters can be recovered from the a_i's via

$$\begin{aligned}
A &= \frac{A_+ + \sqrt{A_+^2 - A_\times^2}}{2} \\
\psi &= \frac{1}{2}\arctan\left(\frac{A_+ a_4 - A_\times a_1}{-(A_\times a_2 + A_+ a_3)}\right) \\
\iota &= \arccos\left(\frac{-A_\times}{A_+ + \sqrt{A_+^2 - A_\times^2}}\right) \\
\varphi_0 &= \arctan\left(\frac{c(A_+ a_4 - A_\times a_1)}{-c(A_+ a_2 + A_\times a_3)}\right)
\end{aligned} \tag{3.40}$$

where

$$\begin{aligned}
A_+ &= \sqrt{(a_1 + a_4)^2 + (a_2 - a_3)^2} \\
&\quad + \sqrt{(a_1 - a_4)^2 + (a_2 + a_3)^2} \\
A_\times &= \sqrt{(a_1 + a_4)^2 + (a_2 - a_3)^2} \\
&\quad - \sqrt{(a_1 - a_4)^2 + (a_2 + a_3)^2} \\
c &= \mathrm{sign}(\sin(2\psi)) \,.
\end{aligned} \tag{3.41}$$

The preceding description of the F-statistic automatically incorporates the two independent LISA channels through the use of the dual-channel noise weighted inner product $(a|b)$. The basic F-statistic can easily be generalized to handle N sources.

Writing $i = 4K + l$, where K labels the source and $l = 1 \rightarrow 4$ labels the four filters for each source, the F-statistic (3.39) keeps the same form as before, but now there are $4N$ 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, and thus easily inverted despite its large size.

CHAPTER 4

DATA ANALYSIS OF A SINGLE COALESCING BINARY SYSTEM

Introduction

This chapter focuses on the response of space-based gravitational wave detectors to waveforms emanating from coalescing binary systems. The bulk of this work was first published in Ref. [23].

ALIA and BBO have been proposed [24, 25] as follow on missions to LISA. Here a study of the capabilities of these observatories is undertaken. Also included is a discussion of how these capabilities relate to the science goals of the missions. These studies indicate that by including a second identical constellation of spacecraft in a 20° Earth-leading orbit, providing a constellation separation of 40°, the parameter resolution will be greatly improved and will go considerably further toward meeting ALIA's main scientific goal of studying IMBHs. This dual constellation configuration will be called the ALIAS. Also addressed is the natural question, given the increases in precision that ALIAS provides relative to ALIA, of what would be gained by adding a second constellation to the LISA mission - a configuration which will be called LISAS. Lastly, the parameter resolution of the BBO is analyzed for NSs, $10M_\odot$, and $100M_\odot$ black holes (BHs), and for the initial deployment phase of the BBO, which will be called BBO Star, containing the two star constellations of the BBO configuration, but not the outrigger constellations. It is shown that parameter resolution of the BBO and BBO Star should be sufficient to detect foreground binary systems, out to and beyond a redshift of $z = 3$.

This chapter is organized as follows: After the executive summary, the basic scientific goals of each of the proposed gravitational wave detectors is briefly covered, followed by a description of how a pair of separated detectors can significantly improve the angular resolution for transient sources. Then results for the ALIA and ALIAS missions for both equal and unequal mass binary systems are given. Next, is a comparison of the LISA and LISAS missions. The final examinations look at missions involving the full BBO mission and a mission involving only the BBO Star constellations.

Chapter Executive Summary

While the survey presented in this chapter is by no means comprehensive, it will help to map out the science that can be done with the ALIA and BBO missions. It will show that ALIAS, which is a modest extension to the ALIA mission, would be able to return a far more accurate census of the IMBH population. In addition the studies will show that a similar extension to LISA would greatly improve its ability to locate the host galaxies of coalescing binaries. On the other hand, a willingness to give up some of the precision astronomy offered by the full BBO, would allow for filling the primary science goal of the BBO with just the first phase of the BBO deployment. The BBO Star configuration could satisfy the primary goal of detecting the GWB, while still providing a detailed binary census.

Space-based Detector Mission Science Goals

It is hoped that LISA will be the first of several efforts to explore the low frequency portion of the gravitational wave spectrum accessible to space borne interferometers. LISA is sensitive to gravitational waves with frequencies ranging between roughly 0.1 mHz and 100 mHz. This bandwith will be a source-rich region of monochromatic binary systems in the galaxy. Also expected to exist in this frequency band are coalescing SMBH binaries with masses above $\sim 10^3 M_\odot$. While such events are thought to be rare, the sensitivity of LISA combined with the power of the gravitational waves created by such an occurrence should allow detections even out to redshifts greater than ten, so the volume of the Universe observable in which such events are detectable by LISA is immense.

The primary mission of ALIA will be detecting IMBHs, with masses in the range $50 - 50,000 M_\odot$. These black holes coalesce at higher frequencies than the SMBHs and produce an inherently weaker gravitational wave. ALIA features a spacecraft configuration that is similar to that of LISA, with arm lengths a tenth the size, a position noise budget 200 times smaller than LISA, and an acceleration noise budget 10 times smaller than LISA. These changes improve the sensitivity at frequencies greater than 10 mHz and provide the ability to detect the weaker IMBH signals. Such detections will provide information on populations, locations, and event rates that could greatly enhance theories of black hole formation and evolution.

The proposed BBO mission will be an extremely sensitive antenna that is designed to detect the GWB left by the Big Bang. According to the standard cosmological picture, the GWB is a relic of the early inflationary period of the Universe. Just as the COsmic Background Explorer (COBE) and the Wilkinson Microwave Anisotropy Probe (WMAP) missions provided information about the Universe around the time of last scattering, the BBO should be able to provide information about the earliest moments in the history of the Universe. The current BBO proposal calls for four LISA-like spacecraft constellations with arm lengths 100 times smaller than LISA, a position noise budget one million times smaller than LISA, and an acceleration noise budget 100 times smaller than LISA. This will push the sensitivity of the detector to higher frequencies ~ 0.1 Hz. Two of the constellations will be centered on a 20° Earth-trailing orbit, rotated 60° with respect to each other in the plane of the constellations (see Figure 2.5). These constellations will be referred to as the star constellations, as the legs of the constellation sketch out a six pointed star. The remaining two constellations are to be placed in an Earth-like orbit 120° ahead and behind the star constellations. These two constellations will be referred to as the outrigger constellations. The purpose of the outrigger constellations is to provide greater angular resolution for foreground sources (see Section 4 for details). The star constellations provide maximum cross-correlation of gravitational wave signals between the two constellations, with minimal correlated noise [14]. The noise in each detector is expected to be independent, so over time the overlap of the noise between the two constellations will tend to average to zero, while the overlap of the signal will grow. The plan is to deploy the BBO in stages, starting with the star constellation, then adding the outrigger constellations at a later date.

Figure 4.1 shows the detector sensitivities for LISA, ALIA, and BBO, a level for the expected extra-galactic background confusion noise [60], and optimally filtered amplitude plots for equal mass binaries in their last year before coalescence at redshift $z = 1$. The squares shown on the amplitude plots denote the frequency of the signal one week before coalescence. The signal from coalescing equal mass binaries at $z = 1$ with masses above $10^1 M_\odot$ will be detectable by both ALIA and BBO. The BBO will also be able to detect lower mass systems such as NS binaries. However, for coalescing binaries, a large portion of the signal strength is due to the rapid inspiral in the last week, and for $10 M_\odot$ binaries at $z = 1$ or greater, much of this power is deposited after the signal has crossed the ALIA sensitivity curve.

Table 4.1 lists the instrument parameters used in this study of the LISA, ALIA and BBO missions. The instrument noise in the A and E channels was modeled by [45]

$$S_n(f) = 8\sin^2(f/2f_*)[\left(2+\cos^2(f/f_*)\right)S_{\rm pos} \\ +2\left(3+2\cos(f/f_*)+\cos(2f/f_*)\right)\frac{S_{\rm accl}}{(2\pi f)^4})] \tag{4.1}$$

where $S_{\rm pos}$ and $S_{\rm accl}$ are the one-way position and acceleration noise contributions.

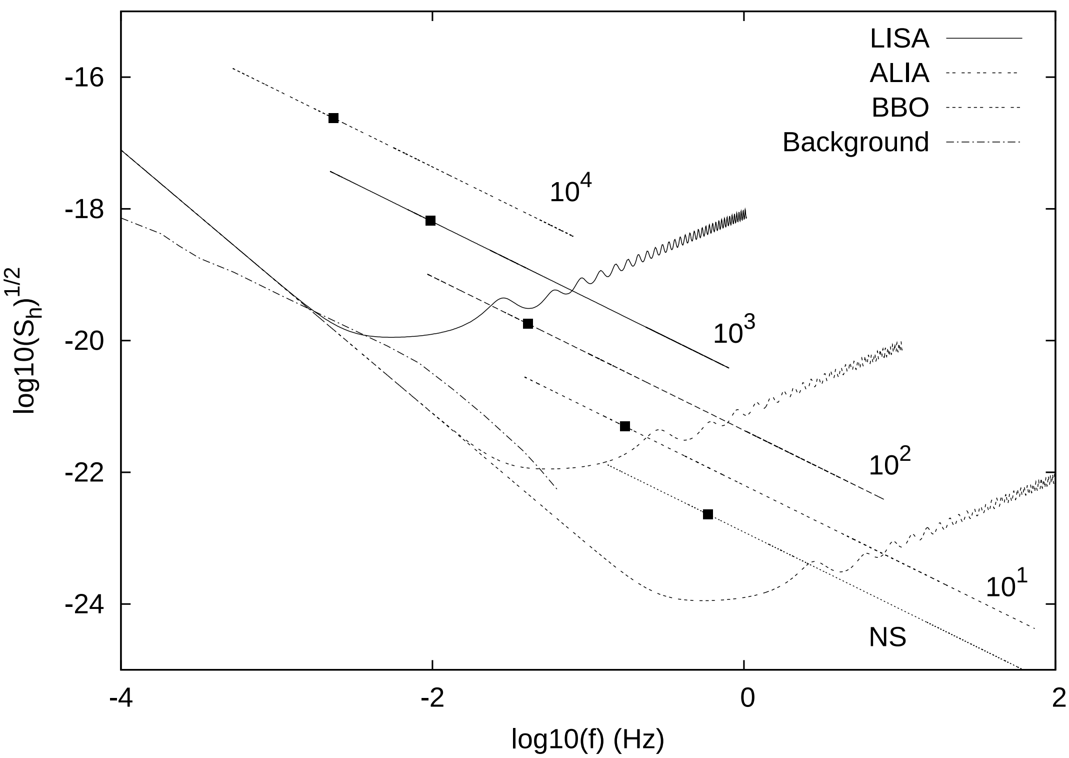

Figure 4.1: Sensitivity curves for LISA, ALIA, and BBO with optimally filtered amplitude plots for equal mass binaries at $z = 1$ in their last year before coalescence. The masses shown are in units of solar masses. The squares denote the frequency one week from coalescence.

Multiple Constellation Detection

Detectors such as LISA and ALIA determine the positions of gravitational wave sources through both amplitude and frequency modulation. The angular resolution

Table 4.1: Instrument Parameters

Parameter	LISA	ALIA	BBO
S_{pos} ($\text{m}^2\,\text{Hz}^{-1}$)	4×10^{-22}	1×10^{-26}	2×10^{-34}
S_{accel} ($\text{m}^2\,\text{s}^{-4}\,\text{Hz}^{-1}$)	9×10^{-30}	9×10^{-32}	9×10^{-34}
L (m)	5×10^{9}	5×10^{8}	5×10^{7}

improves with time due to the accumulation of SNR and the synthesis of a long baseline as the detectors move in their orbit. In contrast, a detector array like the BBO has widely separated elements, and thus has a built-in baseline. Adding a second widely separated constellation to LISA or ALIA would increase the SNR by a factor of $\sqrt{2}$, but the main gain in angular resolution for transient sources would be due to the built-in baseline. The advantage of a multi-element array can be understood from the following toy model. Suppose that there is a gravitational wave of known amplitude and frequency. Neglecting the effects of amplitude modulation, the detector response is given by:

$$h = A\cos(2\pi f(t + R\sin(\theta)\cos(2\pi f_m t + \kappa - \phi))) \tag{4.2}$$

Here f is the source frequency, R is the distance from solar barycenter to the guiding center of the constellation, $f_m = 1/\text{year}$ is the modulation frequency, κ is the azimuthal location (along Earth's orbit) of the guiding center, and θ and ϕ give the source location on the sky.

With this two parameter (θ and ϕ) signal one can analytically derive the uncertainty in the solid angle for a source observed for a time, T_{obs}, by a single constellation

$$\Delta\Omega_{\text{single}} = \frac{2S_n(f)}{(A\pi f R)^2\sin(2\theta)T_{obs}\sqrt{(1-\text{sinc}^2(2\pi f_m T_{obs}))}}\,. \tag{4.3}$$

For small observation times, $\Delta\Omega_{\text{single}}$ scales as T_{obs}^{-2}, while for large observation times it scales as T_{obs}^{-1}.

Turning to the dual detector case, one can simply add together the FIMs for each individual detector. For two constellations separated by an angle $\Delta\kappa$, this yields a solid angle uncertainty of

$$\Delta\Omega_{\rm dual} = \frac{S_n(f)}{(A\pi f R)^2 \sin(2\theta) T_{obs}\sqrt{(1-\mathrm{sinc}^2(2\pi f_m T_{obs})\cos^2(\Delta\kappa))}}\,. \tag{4.4}$$

For small observation times and non-zero $\Delta\kappa$, $\Delta\Omega_{\rm dual}$ scales as T_{obs}^{-1}. In other words, the built-in baseline leads to a much improved angular resolution for short observation times. This is very important for coalescing binaries as most of the SNR accumulates in the final days or weeks prior to merger. Note that if the two constellations are co-located, $\Delta\kappa = 0$, the uncertainty in the solid angle is reduced by a factor of two relative to the single detector case by virtue of the increased SNR. Also, note that this toy model only includes Doppler modulation, thus the symmetry between $\Delta\kappa = 0$ and $\Delta\kappa = \pi$. Including the amplitude modulation breaks this symmetry.

Coalescing Binary Systems

Coalescing binary systems, in contrast to the so-called monochromatic binaries, show large notable increases in their orbital and gravitational wave frequencies during a relatively limited time of observation. In general, the closer a binary system is to coalescence the more rapid its frequency changes. This is one aspect of the gravitational wave chirp, mentioned in Chapter 2. For monochromatic sources, the frequency of the gravitational wave was treated as a fixed parameter of the system, that is not the case for coalescing systems. Also, monochromatic binary systems were fully described by a set of 7 parameters, while coalescing systems, with circular orbits and non-spinning constituent masses, require 9 parameters.

Looking at the nature of the gravitational wave frequency for a coalescing system one can discover how the extra parameters enter into the problem. To first order in the (small) post-Newtonian expansion parameter

$$x \equiv \left(\pi(M_1 + M_2)(1+z)f\right)^{2/3}, \tag{4.5}$$

where z is the redshift of the binary, the frequency evolves according to:

$$f(t) = \left(\frac{5}{t_c - t}\right)^{3/8} \frac{1}{8\pi(\mathcal{M}(1+z))^{5/8}} \left(1 + \frac{1}{2}\left(\frac{743}{336} + \frac{11\mu}{4(M_1 + M_2)}x\right)\right). \tag{4.6}$$

Here $\mu \equiv \frac{M_1 M_2}{M_1 + M_2}$, is the standard reduced mass of the binary system, $\mathcal{M} \equiv \frac{(M_1 M_2)^{3/5}}{(M_1+M_2)^{1/5}}$ is the chirp mass of the binary system, and t_c is the time of coalescence for the binary system. These three parameters will be needed to describe the gravitational wave frequency of a coalescing binary system at any given time, t. The parameter set, $\{\mu,\ \mathcal{M},\ t_c\}$, for coalescing binary systems replace the frequency parameter used to describe monochromatic binaries.

The other aspect of the chirp involves the increasing of the gravitational wave amplitude as the two objects spiraled toward each other. This varying amplitude of the wave is given, to lowest order, by [44]:

$$A(t) = \left(\frac{5}{96}\right)^{1/2} \frac{(\mathcal{M}(1+z))^{5/6}}{\pi^{2/3} D_L} f(t)^{-7/6}, \tag{4.7}$$

where D_L is the luminosity distance to the binary system. The luminosity distance will take the place of the monochromatic amplitude parameter to fill out the set of 9 parameters.

Results

The data shown here is for sources at a redshift of $z = 1$ for ALIA, ALIAS, LISA, and LISAS and sources at redshift $z = 3$ for BBO, and BBO Star. These correspond to luminosity distances of 6.63 Gpc and 25.8 Gpc, respectively, using the best fit WMAP cosmology [61]. The uncertainties in the nine parameters are determined using a FIM approach. Each binary system is observed for the last year before coalescence. Each data point is distilled from 10^5 random samples of θ, ϕ, ι, ψ, and φ_0. For t_c, D_L, $\mathcal{M}$, and μ logarithmic derivatives are used so that the uncertainties listed have been scaled by the value of the parameters. The uncertainty in sky location is simply the root of the solid angle uncertainty ($\sqrt{\Delta\Omega}$). The remaining angular parameters, ι, ψ, and φ_0, have not been scaled. The signals are modeled using a restricted second-Post Newtonian (2PN) approximation [62], whereby the amplitude is kept to Newtonian order while the phase is kept to second order. In other words, only the dominant second harmonic of the orbital frequency is included.

The angular variables were chosen by using a Monte Carlo method. The values for $\cos(\theta)$ and $\cos(\iota)$ were chosen from a random draw on $[-1, 1]$. Values for ϕ, ψ, and φ_0 were each chosen from random draws on $[0, 2\pi]$. The parameters $\mathcal{M}$ and D_L were set for each Monte Carlo run (though they changed between runs). Time to coalescence, t_c, was set to 1 year plus a small offset so that during the year of observation the binary did not reach a relativistic regime that would not be properly modeled by the 2PN approximations used.

The mean SNRs quoted in this chapter are calculated by taking the square root of the average of the squares of the individual SNRs. For this analysis, positive detection will be restricted to SNRs above 5.

Results for ALIA

Table 4.2 summarizes the medians and means of parameter uncertainties for detections by ALIA of equal mass binaries with masses of $10^2 M_\odot$, $10^3 M_\odot$, $10^4 M_\odot$, and $10^5 M_\odot$. The SNRs for this range of masses shows that ALIA should get positive detection of $99^{+}\%$ of coalescing IMBHs located at $z = 1$ or closer, which would provide good information on the coalescence rates. The great precision in the measurement of $\mathcal{M}$ and μ will provide a clear picture of the constituent masses of the binary systems. Furthermore, the sub-degree precision in the sky location, combined with luminosity distances known to a few percent, will facilitate the construction of a three dimensional distribution of IMBHs with which to test theoretical predictions (see Ref. [63] for a related discussion concering LISA and supermassive black holes).

The uncertainties in t_c a month from coalescence (using 11 months of data) will be on the order of a few minutes. This will provide warning time for ground-based gravitational wave detectors, as well as other systems (telescopes, neutrino detectors, etc.) to gather as much and as varied data as possible about the coalescence.

The shapes of the histograms shown in Figure 4.2, which are from the $10^3 M_\odot$ data, are representative of the histograms for each of the detectors covered in this work. Note that for ι, ψ, and φ_0, the tails of the histograms run far beyond the range of the plots shown, raising the values of the means considerably above their respective median values. For example, while the median value of the uncertainty in φ_0 for equal mass binaries with masses $10^2 M_\odot$ is $12.2°$ its mean value is $2185°$, which is well beyond the $[0, 2\pi]$ range of φ_0. Uncertainty ranges that are larger than the possible range of the parameter show that the parameter is indeterminate. When uncertainty ranges exceed the possible range of a parameter, one may drop that parameter from the FIM analysis. For this analysis parameters were not discarded. As an example of how this affects the analysis, consider the $10^2 M_\odot$ study where the mean value of the uncertainties that lie in the $[0, 2\pi]$ range for φ_0 would be $36.3°$, while its median would be $9.79°$. However, for 16.9% of the binaries, the φ_0 parameter would be indeterminate.

Figure 4.3 shows the SNR histograms for $10 M_\odot$ equal mass binaries for ALIA (and

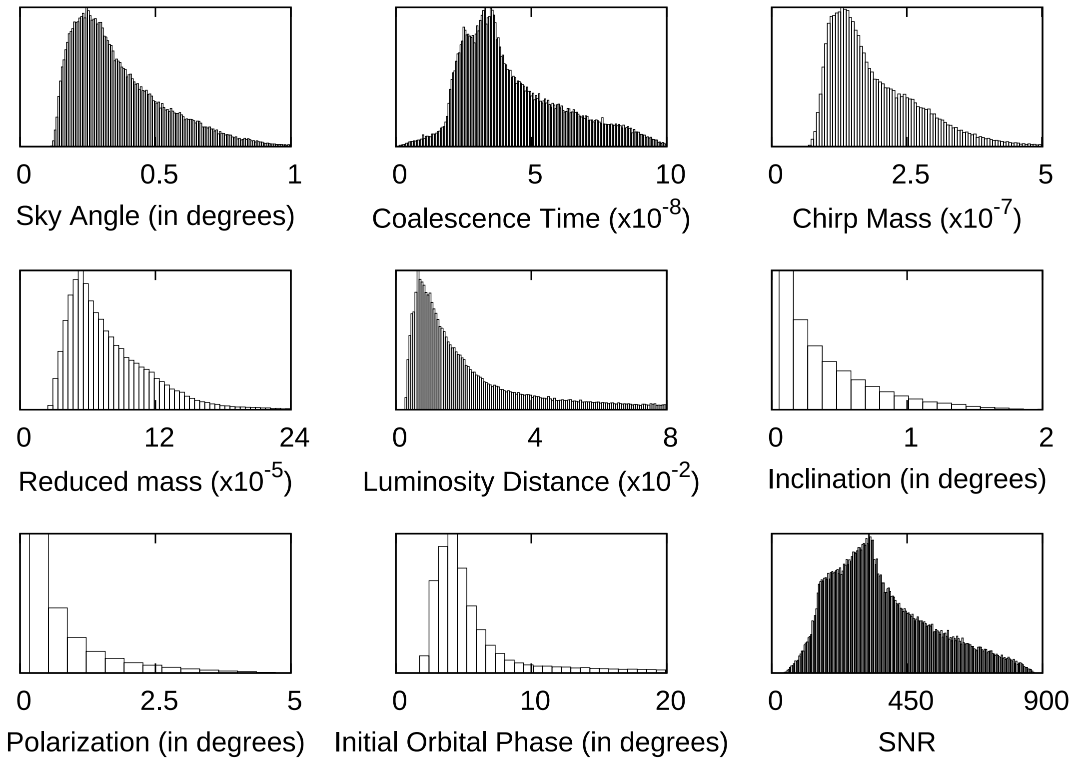

Figure 4.2: Histograms of the parameter uncertainties and SNR, for equal mass binaries of $10^3 M_\odot$ at $z = 1$, as detected by ALIA.

Table 4.2: SNR and parameter uncertainties for ALIA with sources at $z = 1$.

	$10^2 M_\odot$		$10^3 M_\odot$		$10^4 M_\odot$		$10^5 M_\odot$	
	median	mean	median	mean	median	mean	median	mean
SNR	71.7	76.2	466	567	2937	3641	6481	7970
Sky Loc.	0.169°	0.201°	0.234°	0.264°	0.382°	0.417°	0.518°	0.546°
$\ln(t_c) \times 10^{-8}$	1.38	1.58	2.74	3.11	4.41	5.16	5.38	6.47
$\ln(\mathcal{M}) \times 10^{-8}$	9.56	11.0	12.1	13.9	31.3	35.7	136	152
$\ln(\mu) \times 10^{-5}$	46.8	53.6	5.10	5.86	1.76	2.03	2.66	3.03
$\ln(D_L) \times 10^{-2}$	6.25	32.7	1.34	6.44	1.08	3.50	1.15	3.08
ι	3.99°	125°	0.762°	22.4°	0.513°	8.00°	0.515°	5.66°
ψ	5.44°	1092°	1.08°	176°	0.840°	42.3°	0.931°	24.7°
φ_0	12.2°	2185°	4.15°	352°	2.79°	85.4°	2.51°	50.0°

ALIAS). As was expected from Figure 4.1, the SNR values for ALIA are low, with nearly 60% too low for a positive detection. Of note, though, is that the SNR scales

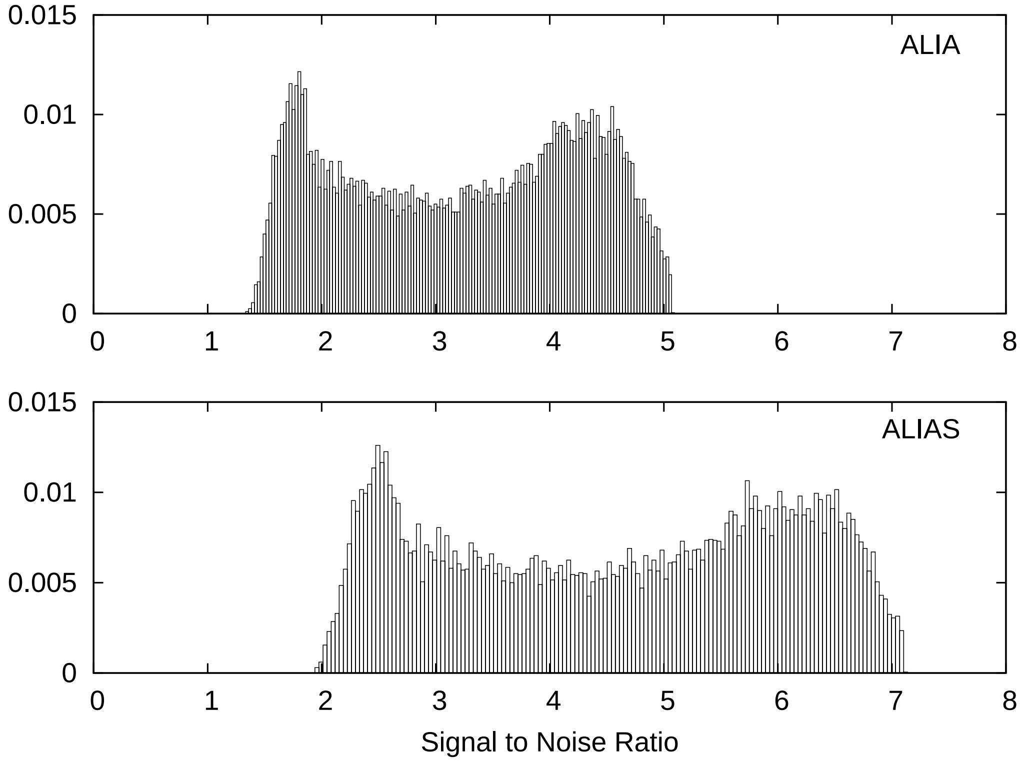

Figure 4.3: SNR histogram for ALIA and ALIAS from $10M_{\odot}$ equal mass binaries at $z = 1$.

roughly as the luminosity distance. This suggests that ALIA should be able to detect nearly all $10M_{\odot}$ binaries with luminosity distances less than 2 Gpc. For binaries with masses beyond $15M_{\odot}$ the low end of the range of SNRs is above 5, meaning that equal mass binaries with masses in the range of IMBHs should be detectable out beyond $z = 1$.

Results for ALIAS

The main purpose of ALIA is gathering information about IMBHs; these data show it is capable of doing this with some success. A more accurate IMBH census could be derived from a dual constellation version of ALIA that is called ALIA in Stereo or ALIAS. Each component of the ALIAS constellation would be offset from the Earth by 20°, one in an Earth-trailing and the other in an Earth-leading orbit, giving a 40° separation in order to provide increased parameter resolution in the IMBH range.

Table 4.3 summarizes the medians and means of the parameter resolutions that could be achieved by the ALIAS mission. Results are given for equal mass binaries with masses of $10^2 M_\odot$, $10^3 M_\odot$, $10^4 M_\odot$, and $10^5 M_\odot$. As expected, the mean SNR increases by $\sqrt{2}$ relative to ALIA. The improvements in parameter resolution are, however, considerably larger. At the upper end of the IMBH mass range the angular resolution improves by a factor of ~ 90 and the luminosity distance resolution improves by a factor of ~ 23.

For masses below $\sim 10^2 M_\odot$, ALIAS provides roughly a factor of $\sqrt{2}$ increase in the median values of each of the parameter uncertainties. This can be seen in the $10^2 M_\odot$ data in Tables 4.2 and 4.3, as well as in Figure 4.4 and Figure 4.5, which plot the uncertainty in the sky location and luminosity distance, respectively, against the chirp mass of the binaries. The trend shown in these figures holds true for D_L, ι, and ψ in that mass range, while the increases in t_c, μ, φ_0, and $\mathcal{M}$ are slightly larger, ~ 1.55 better than the values for ALIA. However, the median SNRs are below 5 for equal mass binaries with chirp masses below $\sim 8 M_\odot$ for ALIAS (below $\sim 10 M_\odot$ for ALIA).

Figure 4.4 also shows that the benefits of a dual constellation becomes even more significant when the final chirp of the binary occurs above the detector noise (see Figure 4.1). The main improvement is in the angular resolution, but the decreased covariances between the sky location and other parameters lead to improved measurements of t_c, D_L, ι, and ψ at the upper end of the IMBH mass range. The resolution of these parameters improves (relative to ALIA) by a factors of 33, 23, 18, and 21, respectively. The parameters $\mathcal{M}$, μ, and φ_0 only see a factor of ~ 2 improvement in resolution.

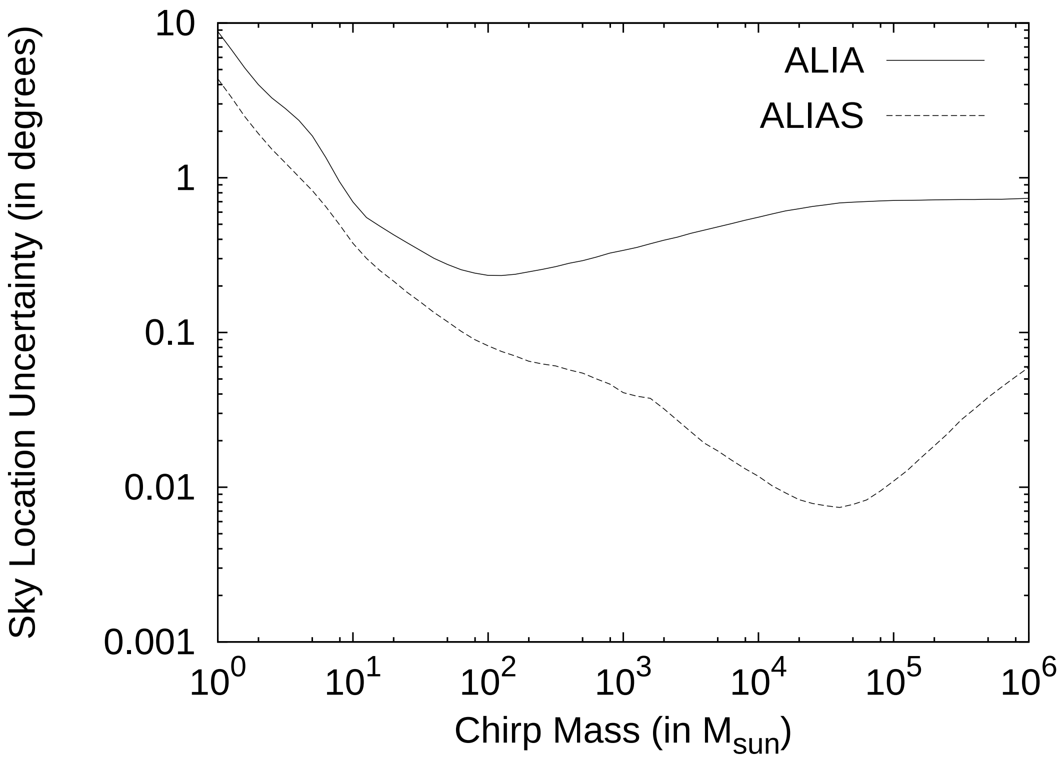

Figure 4.4: Uncertainty in the sky location for both ALIA and ALIAS plotted against binary chirp mass. Binary system constituents in this plot have equal masses and are located at $z = 1$.

Comparing Figure 4.5 with Figure 4.4, one can see that the increased resolution in the luminosity distance corresponds with the increased precision in sky location due to their large covariance. A similar effect was seen in the work of Hughes and Holz [64], where the addition of electromagnetic information to fix the sky location of a BH merger resulted in a marked increase in resolution of the luminosity distance. In the present case, the degeneracy is broken by the improved angular resolution afforded by the baseline between the ALIAS detectors. These improvements in sky location and luminosity distance resolution will provide an even more detailed three dimensional distribution of the IMBHs than that provided by ALIA. Increasing the $3D$ resolution of the distribution of IMBHs by roughly a factor of 6 on the low end of the mass range ($50M_{\odot}$) up to $\sim 200,000$ on the high end of the mass range

$(50,000 M_\odot)$, would provide much more detailed information compared to ALIA, with which to test theoretical predictions about IMBHs.

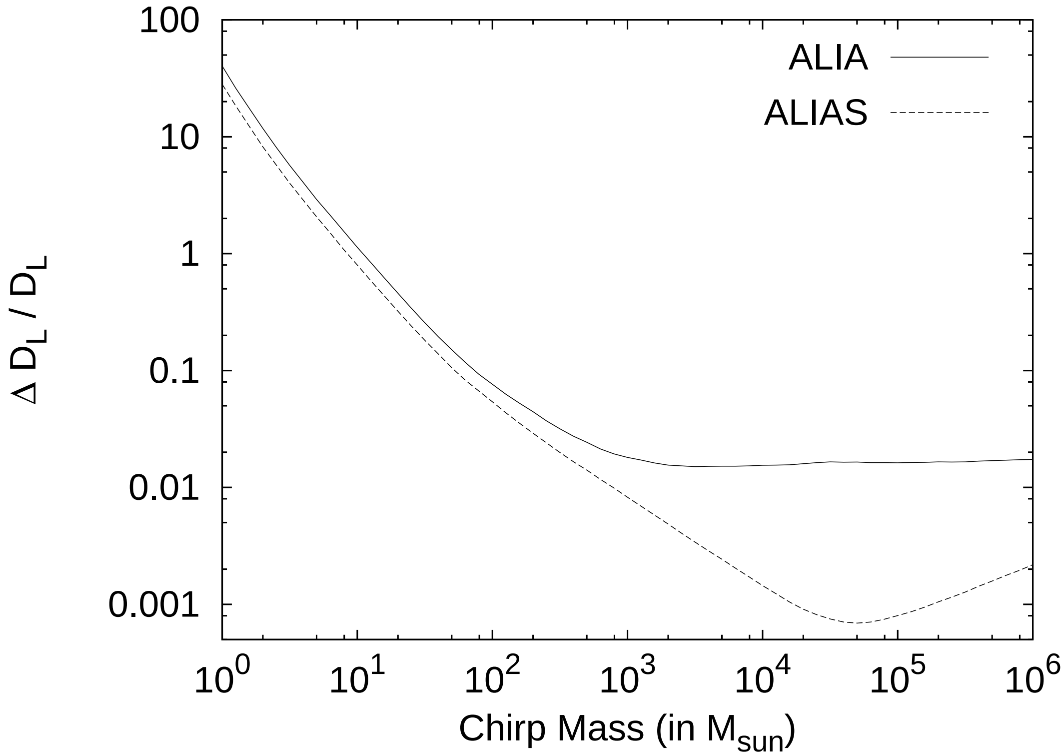

Figure 4.5: Uncertainty in the luminosity distance of both ALIA and ALIAS plotted against binary chirp mass. Binary system constituents in this plot have equal masses and are located at $z = 1$.

Table 4.3: SNR and parameter uncertainties for ALIAS with sources at at $z = 1$.

	$10^2 M_\odot$		$10^3 M_\odot$		$10^4 M_\odot$		$10^5 M_\odot$	
	median	mean	median	mean	median	mean	median	mean
SNR	102	108	688	802	4402	5164	9630	11371
Sky Loc.	0.0620°	0.0769°	0.0284°	0.0330°	0.00786°	0.00905°	0.00757°	0.00846°
$\ln(t_c) \times 10^{-9}$	7.84	9.01	10.6	11.9	3.33	4.42	2.19	2.82
$\ln(\mathcal{M}) \times 10^{-8}$	6.04	6.86	6.64	7.61	18.3	21.1	80.8	93.0
$\ln(\mu) \times 10^{-5}$	23.0	26.2	2.46	2.8	1.05	1.19	1.7	1.94
$\ln(D_L) \times 10^{-3}$	43.7	231	6.67	34.9	1.15	5.88	0.553	2.78
ι	2.80°	89.7°	0.413°	13.1°	0.0678°	2.17°	0.0324°	1.03°
ψ	3.86°	794°	0.575°	112°	0.105°	18.2°	0.0501°	8.61°
φ_0	7.96°	1589°	2.27°	226°	1.34°	37.1°	1.23°	18.0°

Figure 4.3 shows the SNR histogram for $10M_\odot$ equal mass binaries for ALIA and ALIAS. While more than two-thirds of the sources at $z = 1$ would be positively detected, nearly one-third would be missed. However, these results suggest that ALIAS will be able to detect nearly all of the $10M_\odot$ binaries with luminosity distances less than 3 Gpc. This is roughly in keeping with the general trend of ALIAS's capabilities compared to ALIA.

As was seen in Figure 4.1 for masses below $10M_\odot$, the source's signal lies below the sensitivity curve for the week prior to coalescence. For larger masses, more of the binary's final chirp will be detectable. This is why the angular resolution of ALIAS becomes so much better than ALIA's at the upper end of the IMBH mass range. For masses above $10^5M_\odot$ the final chirp occurs before the sweet spot of the sensitivity curve, which diminishes ALIAS's advantage over ALIA.

Results for ALIA and ALIAS for Unequal Mass Binaries

Equal mass binaries are the easiest to study, but they can give an overly optimistic picture of the instrument capabilities as they yield the smallest parameter uncertainties. To study this bias and provide a more realistic picture of the capabilities of ALIA and ALIAS, the case is considered where the masses of the components in the binary are drawn from logarithmic distributions in the range $1 - 10^8M_\odot$ (i.e. a distinct mass was drawn for each set of angular parameters, with the mass of the second binary constituent then determined by the current value of the chirp mass). The results are still presented as a function of chirp mass, but the mass ratios reflect the underlying mass distribution.

Figure 4.6 shows the plot of sky location uncertainty against chirp mass. The uncertainties in sky location for masses above $10^2M_\odot$ are larger than those shown in Figure 4.4. The increase in sky angle uncertainty relative to the equal mass case was a factor of ~ 2 for ALIA. Similarly, other parameters show a factor of ~ 3 increase in their uncertainties relative to the equal mass study.

However, ALIAS still maintains an advantage in locating binaries. In fact, for chirp masses below $\sim 500 M_\odot$ the increase in precision for ALIAS over ALIA in locating unequal binaries is nearly the same as it was for equal mass binaries, a factor of roughly $\sqrt{2}$. At higher chirp masses one sees less of a difference in angular resolution than in the equal mass study (factors of ~ 9 compared to ~ 90). This trend holds for t_c, D_L, ι, and ψ, which show maximum increases in precision ~ 4 for ALIAS over ALIA, while $\mathcal{M}$, μ, and φ_0 show maximum increases in precision ~ 2 for ALIAS over ALIA (as they did for equal mass binaries). Thus the increased resolution in the $3D$ distribution of IMBHs for ALIAS over ALIA for unequal mass binaries ranges from a factor of ~ 6 up to ~ 300.

Results for LISA and LISAS

Figure 4.7 compares the sky location uncertainties for the LISA mission to the LISAS mission. As was seen with ALIA and ALIAS, the addition of a second constellation to LISA provides a marked increase in parameter resolution. The increases in parameter resolution for LISAS over LISA are similar to those found between ALIAS and ALIA. The angular resolution showed a maximum improvement of ~ 25 just above $10^5 M_\odot$. Also, a lesser benefit than that shown in Figure 4.7 occurs in the t_c, D_L, ι, and ψ parameters (with factors of maximum increase ~ 12 just above $10^5 M_\odot$), while $\mathcal{M}$, μ, and φ_0 show only a modest maximum improvement by a factor of ~ 2. These results are in the same range as those found by Seto [65] for sources detected by LISA that have undergone strong gravitational lensing. In that case the extended baseline was provided by the time delay in the arrival of the signals, which effectively turned a single LISA detector into a LISAS system. Seto also gave brief consideration to the performance of a dual LISA mission, and found the ratio of angular resolution

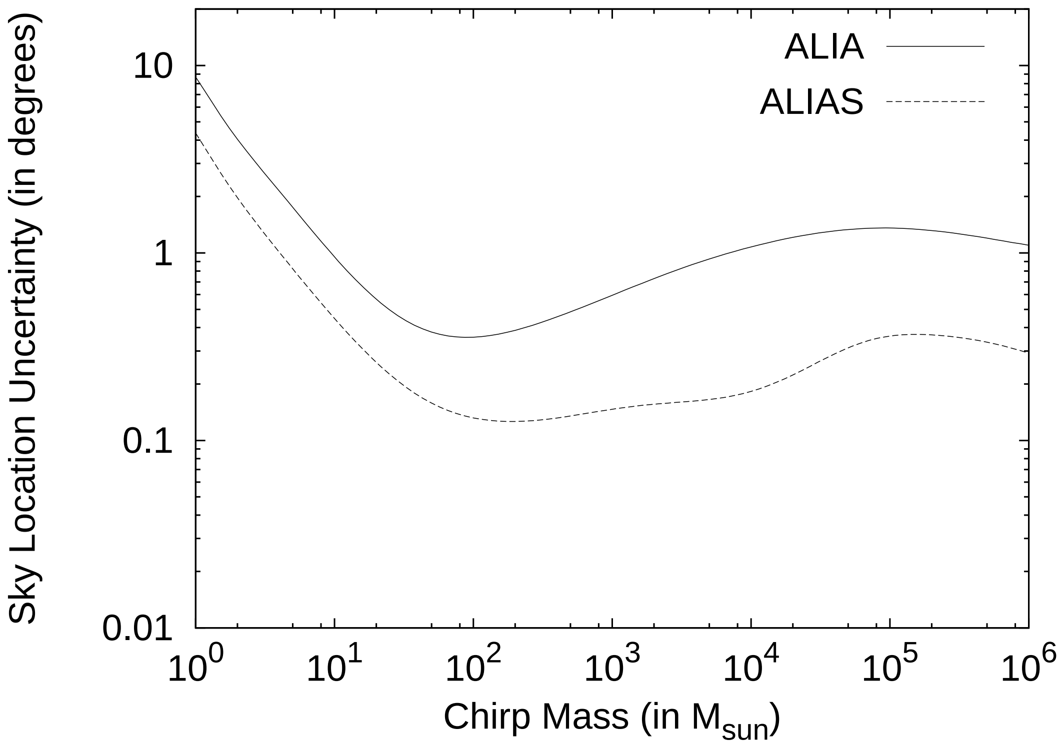

Figure 4.6: Uncertainty in the sky location for both ALIA and ALIAS plotted against binary chirp mass. Binary system constituents in this plot have unequal masses and are located at $z = 1$.

between LISAS and LISA to be larger than that seen in these simulations. The discrepancy can be traced to Seto including a large galactic confusion background which limits the time that the sources are in-band, thus magnifying the advantage of the dual configuration.

The shift in the placement of the minimum uncertainties, compared to ALIA and ALIAS, can be understood from Figure 4.1. The signal from a $10^3 M_\odot$ equal mass binary is passing below the LISA sensitivity curve in the last week before coalescence, and before the final chirp of the binary. This is evident in Figure 4.7, where the angular resolution of the single and dual configurations is similar for masses below $10^3 M_\odot$. The sweet spot of the LISA sensitivity curve is at ~ 5 mHz, which is where $10^5 M_\odot$ binaries experience their final chirp. Above $10^5 M_\odot$, less of the final chirp

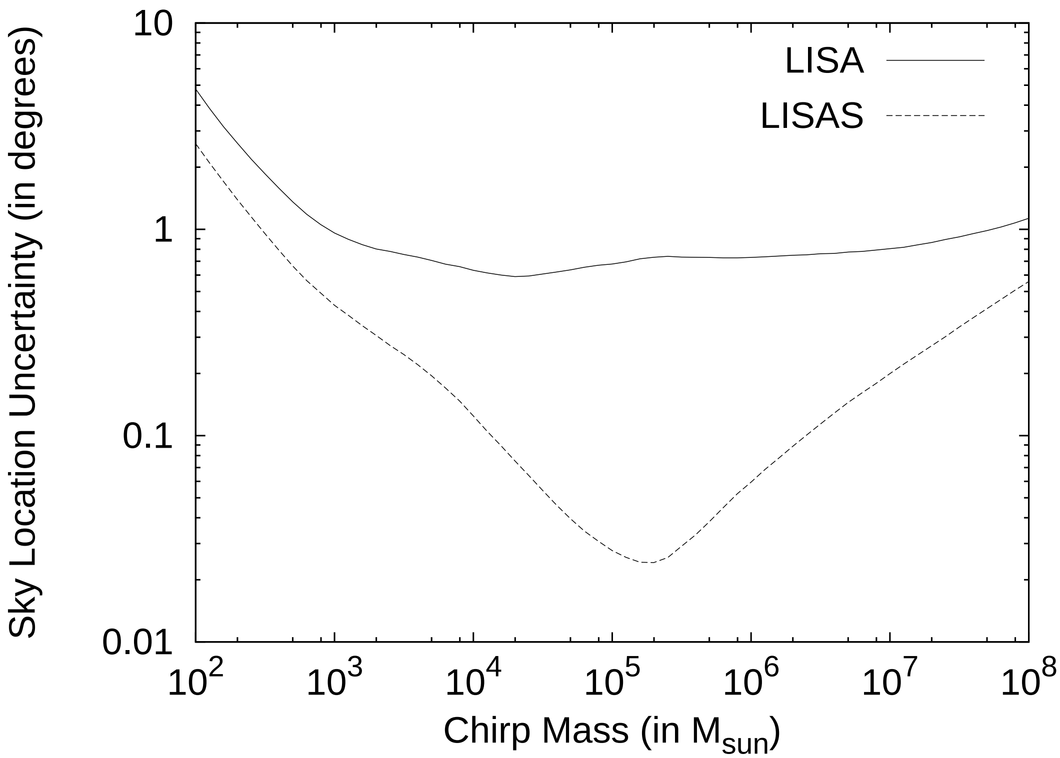

Figure 4.7: Uncertainty in the sky location for both LISA and LISAS plotted against binary chirp mass. Binary system constituents in this plot have equal masses and are located at $z = 1$.

occurs near the sweet spot, and the difference in the angular resolution diminishes. Figure 4.8 shows how this shift corresponds precisely to the shift in the maximum SNR of the LISA and LISAS missions. The plot shows how the increase in the mean SNR (i.e. $\sqrt{2}$) is uniform across the range of chirp masses. Also, it shows again how the resolution in D_L improves with the addition of more information in the form of increased angular resolution.

Results for BBO

Table 4.4 summarizes the medians and means of the parameter uncertainties for detections of equal mass binaries with masses of $1.4M_\odot$, $10M_\odot$, and $10^2M_\odot$ for the BBO. Histograms for these data have the same general shapes as those shown in

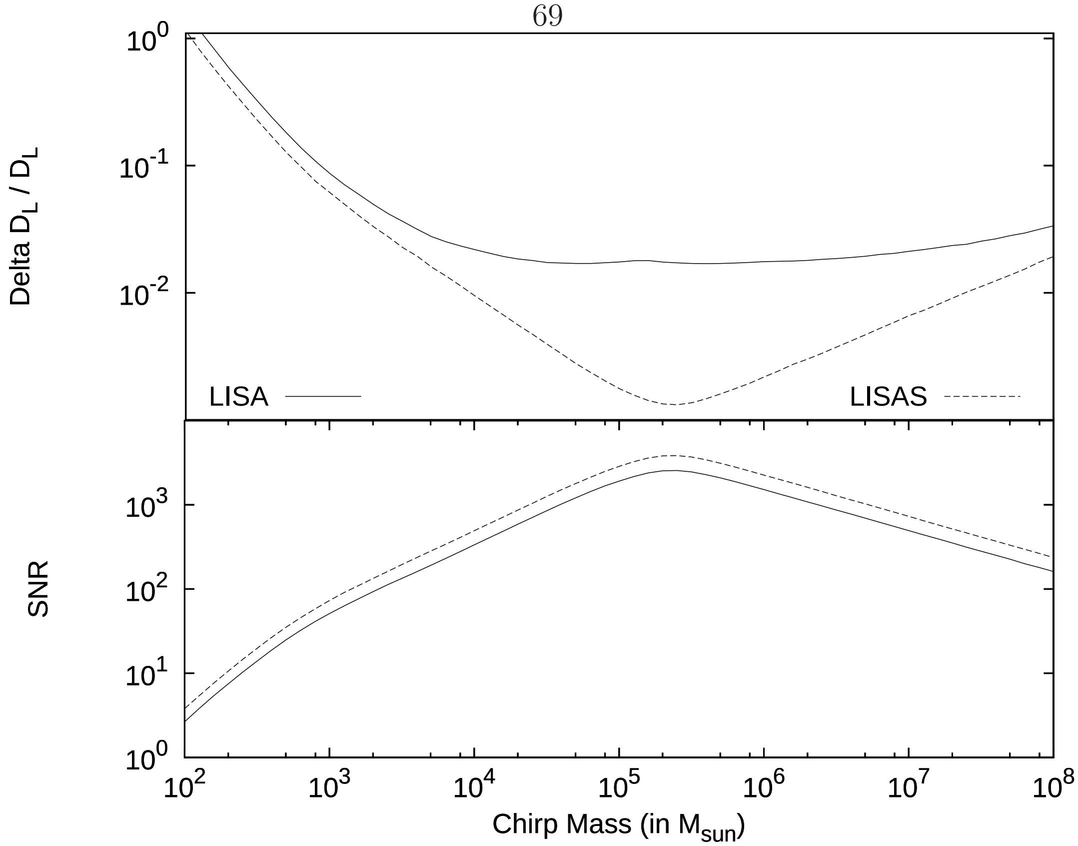

Figure 4.8: Uncertainty in the luminosity distance and the SNR for both LISA and LISAS plotted against binary chirp mass. Binary system constituents in this plot have equal masses and are located at $z = 1$.

Figure 4.2 and so are not shown.

As can be seen from the data, the BBO is an extremely sensitive detector. The SNRs show that if the BBO meets design specifications there will be positive detection of coalescing stellar mass binaries out to (indeed well beyond) a redshift of $z = 3$. Similarly, the BBO will provide a precise picture of coalescence rates, including how these rates relate to luminosity distance. With a combination of sky location and luminosity distance the BBO should be able to pick out the host galaxy of the majority of coalescing binaries. Data taken by the BBO up to a month before coalescence (using 11 months of data) will determine t_c to within seconds, again providing ample warning time for other detectors to gather data on the coalescence.

Table 4.4: SNR and parameter uncertainties for BBO with sources at $z = 3$.

	$1.4M_\odot$		$10M_\odot$		$10^2M_\odot$	
	median	mean	median	mean	median	mean
SNR	156	162	809	843	4958	5189
Sky Loc.	3.31"	4.49"	0.673"	0.920"	0.176"	0.246"
$\ln(t_c) \times 10^{-12}$	179	207	32.1	37.2	8.09	9.28
$\ln(\mathcal{M}) \times 10^{-9}$	63.1	72.7	7.02	8.09	30.4	35.0
$\ln(\mu) \times 10^{-6}$	173	200	16.3	18.8	4.97	5.70
$\ln(D_L) \times 10^{-3}$	28.9	150	5.58	29.1	0.927	4.73
ι	1.86°	57.3°	0.360°	11.1°	0.0592°	1.78°
ψ	2.54°	497°	0.492°	96.1°	0.0806°	15.2°
φ_0	7.71°	996°	3.93°	194°	2.82°	32.3°

Results for BBO Star

As can be seen in Figure 4.1, the final chirp of solar mass or NS coalescing binaries occurs below the noise level of the BBO. Thus the outrigger constellations do not provide an extended baseline for the last few days of NS chirp. Positive detection of NS binaries at $z = 3$ can be accomplished with the initial deployment of the two constellations that make up the star, while still providing the cross-correlating needed to detect the GWB.

Table 4.5 summarizes the medians and means of the parameter uncertainties for the detection of equal mass binaries with masses of $1.4M_\odot$, $10M_\odot$, and $10^2M_\odot$ for the BBO Star. Histograms for these data have the same general shapes as those shown in Figure 4.2 and so are not shown.

Figure 4.9 plots the uncertainty in sky location against chirp mass for the BBO and BBO Star. As can be seen, the effect of the outrigger constellations is significant, providing for increased parameter resolution for the full BBO throughout the range of chirp masses shown, with a maximum increase of ~ 3700 around $200M_\odot$. Similar to the ALIA/ALIAS comparison, this extra precision increases until all of the final chirp lies above the sensitivity curve (see Figure 4.1), and begins to decrease as less and less of the final chirp occurs in the frequency range of the sweet spot of the sensitivity curve (which for BBO is $\sim$ 1 mHz). The improvement in the mean SNR for BBO over BBO Star is the expected $\sqrt{2}$. Also, the improvement in the resolution of t_c (~ 3550) is comparable to that seen in the angular resolution. More modest increases occur for D_L, ι, and ψ, which increase by factors of 52, 35, and 47, respectively. Only slight (~ 2.5)increases are seen for $\mathcal{M}$, μ, and φ_0.

The SNRs from BBO Star are sufficient for positive detection of binaries with constituents less than one solar mass out to, and beyond, $z = 3$. While the full BBO offers considerable advantages for doing precision gravitational wave astronomy, BBO Star could fulfill the main science objective of detecting the cosmic gravitational wave background while still providing useful information about binary populations.

Table 4.5: SNR and parameter uncertainties for BBO star constellations with sources at $z = 3$.

	$1.4M_\odot$		$10M_\odot$		$10^2M_\odot$	
	median	mean	median	mean	median	mean
SNR	96.9	114	484	598	3022	3697
Sky Loc.	82.6”	107”	214”	256”	516”	596”
$\ln(t_c) \times 10^{-9}$	2.98	3.43	7.52	8.60	19.8	22.3
$\ln(\mathcal{M}) \times 10^{-8}$	13.1	15.1	2.00	2.29	6.95	7.92
$\ln(\mu) \times 10^{-5}$	3.70	4.28	3.70	4.31	1.23	1.53
$\ln(D_L) \times 10^{-3}$	49.9	257	12.5	68.0	6.51	31.9
ι	3.02°	95.8°	0.767°	22.9°	0.310°	7.43°
ψ	4.07°	815°	1.04°	176°	0.529°	40.6°
φ_0	14.1°	1633°	9.56°	357°	6.99°	84.9°

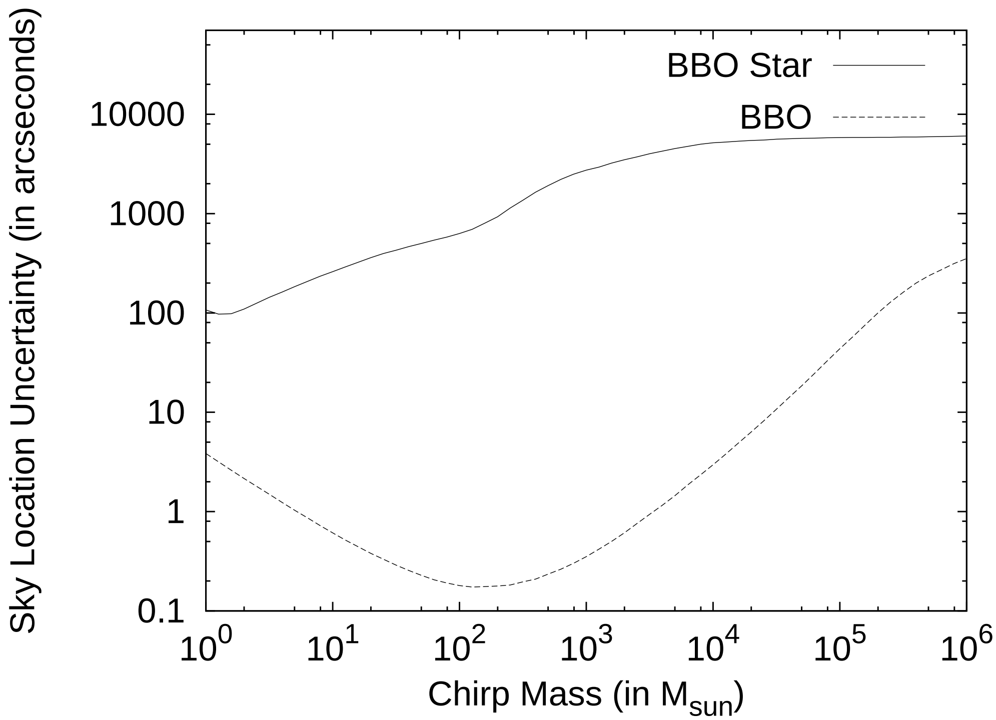

Figure 4.9: Uncertainty in sky location versus chirp mass for the standard BBO design and the star constellations of the BBO. Binary system constituents in this plot have equal masses and are located at $z = 3$.

CHAPTER 5

DATA ANALYSIS OF MULTIPLE MONOCHROMATIC SOURCES

Introduction

This chapter focuses on how the presence of multiple monochromatic sources affects the resolution of LISA. The bulk of this work was first published in Ref. [66].

These studies show that the parameter estimation uncertainties grow exponentially with the number, N, of overlapping sources. This should be contrasted to the $\sqrt{N}$ increase one would predict if the other sources were treated as stationary, Gaussian noise (SGN). The degredation in resolution was found to be nearly uniform across the seven parameters that LISA will measure for a galactic binary system. As one might expect, the parameter uncertainties are a strong function of the signal cross-correlations. It is found that as the two signals become more correlated, the parameter estimation uncertainties grow at a rate that is faster than exponential. Importantly, it is shown that the parameter uncertainties decrease rapidly as the observation time T is increased - far more rapidly than the usual $1/\sqrt{T}$ improvement one expects when competing with SGN.

In the low frequency portion of the gravitational wave spectrum there are expected to be many thousands of compact galactic binaries. The problem of source confusion will be most pronounced below ~ 2 mHz, where it is expected that many tens to tens of thousands of galactic binaries will have signals that overlap in each frequency bin [67]. Thus, attention is focused on low frequency galactic binaries that are close to one another in frequency. A natural question to ask is how the resolution of the source parameters, such as sky location, amplitude and binary orientation, of these binaries will be affected by the presence of other binary systems in the LISA data stream. Estimates of the parameter resolution have been given for compact galactic binaries [44, 68, 69], supermassive black hole binaries [23, 44, 68, 70, 71, 72], and extreme mass ratio inspirals [73]. These studies focused on the problem of identifying one source at a time, and did not address the problem of source confusion. The large signal from the galactic population of close white dwarf binaries was not ignored, but it was treated as an additional source of SGN to be added to the instrument noise. The study presented in this chapter shows that this is not a good approximation.

The chapter is organized as follows: After the executive summary is an extension to the FIM approach detailed in Chapter 3 to N sources. Next is a look at the how parameter resolution and source confusion are affected by the observation time. Then is an analysis of the effect of the separation in frequency of the nearest neighbor binary systems, followed by a study of the relationship between the amount of signal overlap and level of source confusion. Then there is an exploration of how source confusion changes with the frequency of the binary systems. To conclude the chapter a brief information theory perspective is presented.

Chapter Executive Summary

The studies presented in this chapter will show that source confusion is going to significantly impact LISA's ability to resolve individual galactic binaries. It is found that source confusion affects parameter estimation fairly uniformly across the seven paramters that describe a monochromatic galactic binary.

The two most significant findings show that source confusion grows exponentially with the number of correlated sources, and that source confusion decreases rapidly with time of observation.

Source confusion will be a significant problem in the frequency range between

0.01 and 3 mHz, where it is estimated that there are upwards of 10^7 galactic binary systems. The decrease in the parameter uncertainties with time of observation is far greater than the usual $1/\sqrt{T_{obs}}$ decay that occurs when competing against SGN. The fast improvement will be shown in this chapter to be due to the $1/T_{obs}$ decrease in the source cross-correlation, and the sensitive dependence of the parameter uncertainties on the degree of cross-correlation.

These findings suggest that earlier work that treated the galactic background as SGN should be revisited, and that every effort should be made to extend the LISA mission lifetime beyond its nominal 3 year duration.

Multiple Sources

The FIM approach to estimating parameter uncertainties given in Chapter 3 was limited to a single source. The approach can be easily generalized to multiple sources. It is necessary to understand the effect of multiple, correlated sources as there will be a large number of such sources on the low end of the LISA frequency band. For N circular, monochromatic binary systems, the parameter space is $7N$ dimensional, and the parameters are combined in the parameter vector

$$\vec{\Lambda} = (\vec{\lambda}^1, \vec{\lambda}^2, \ldots, \vec{\lambda}^N. \tag{5.1}$$

The notation λ^A refers to all 7 parameters of the A^{th} binary. The total signal $S = S_I$ or $S = S_{II}$ will be the sum of the signals for the individual binary systems,

$$S = S^1 + S^2 + S^3 \ldots + S^N \,, \tag{5.2}$$

where S^A is the signal due to the A^{th} binary. Since $\partial S^A / \partial \lambda_B = 0$ for $A \neq B$, the full

FIM has a simple structure. Defining

$$\Gamma_{AB} = \sum_{\alpha=I,II} \langle \frac{\partial S^A_\alpha}{\partial \Lambda_A} | \frac{\partial S^B_\alpha}{\partial \Lambda_B} \rangle \,, \tag{5.3}$$

each diagonal block Γ_{AA} is the usual 7×7 FIM for the $A^{\rm th}$ source, while the off-diagonal blocks Γ_{AB} describe how the parameter estimation for source A is influenced by the presence of source B and vice-versa. If all the sources are un-correlated (for example, they might be widely spaced in frequency), then the off-diagonal blocks will be zero, and the full FIM will be block diagonal. Upon inverting to get the variance-covariance matrix the block diagonal structure will be preserved, with each 7×7 block equal to the inverse of the corresponding single source FIM. However, if the sources are overlapping, the off-diagonal blocks will be non-zero, which will affect the values of the diagonal elements in the full variance-covariance matrix. The number of off-diagonal elements grows quadratically with the number of sources, while the number of diagonal elements grows linearly with the number of sources. Thus, it is anticipated that the ability to resolve a particular source's parameters will degrade rapidly as the number of correlated sources increases. The volume of the $n\sigma$ uncertainty ellipsoid in the D dimensional parameter space is

$$V_D = \frac{\pi^{D/2} n^d}{\Gamma(D/2+1)} \sqrt{\det C_{ij}} \,, \tag{5.4}$$

where the Γ in the denominator is the Gamma-function.

The parameter uncertainties are also anticipated to depend strongly on the degree of correlation between sources. The correlation between two signals S^A and S^B is defined by:

$$\kappa_{AB} = \frac{(S^A|S^B)}{(S^A|S^A)^{1/2}(S^B|S^B)^{1/2}} \tag{5.5}$$

In analyzing the results, two similar approaches will be used to obtain a quantitative measure of the increase in the parameter estimation uncertainties. First is a global comparison of the uncertainties. The uncertainties form an ellipsoid in the parameter space whose volume is given by the determinant of the variance-covariance matrix. One measure of the uncertainty increase due to the correlation of the binary systems is given by a ratio of the geometric mean of the uncertainties (GMUR),

$$\mathrm{GMUR} \equiv \left(\frac{\det \Gamma}{\prod_A \det \Gamma_{AA}} \right)^{1/7N} . \tag{5.6}$$

The GMUR describes the mean increase in the parameter uncertainties due to source confusion.

A second measure of the uncertainty increase is a parameter by parameter comparison, looking at the ratio of the parameter uncertainty (PUR) between the uncertainty of a particular parameter calculated in the $7N$-dimensional variance-covariance matrix, $\Delta\Lambda^A_x$, and the uncertainty of the same parameter calculated in the isolated binary's variance-covariance matrix, $\Delta\lambda^A_x$.

$$\mathrm{PUR} \equiv \frac{\Delta\Lambda^i_x}{\Delta\lambda^i_x}. \tag{5.7}$$

This definition for the PUR is independent of the amplitudes of the sources, as can be seen in the simple case of two sources, each described by a parameter α. The PUR for α_1 is given by:

$$\mathrm{PUR}_{\alpha_1} = \frac{\Delta\Lambda_{\alpha_1}}{\Delta\lambda_{\alpha_1}} = \frac{1}{\sqrt{1 - \Sigma^2_{\alpha_1\alpha_2}}} , \tag{5.8}$$

where

$$\Sigma_{\alpha_1\alpha_2} \equiv \frac{\Gamma_{\alpha_1\alpha_2}}{\sqrt{\Gamma_{\alpha_1\alpha_1}\Gamma_{\alpha_2\alpha_2}}}. \tag{5.9}$$

The independence of the degree of confusion on the source amplitudes would seem to imply that an arbitarily weak source can affect parameter estimation to the same degree as an arbitarily strong source. Clearly, this makes no sense in the limit that the amplitude of the weak source goes to zero. The resolution to this apparent paradox is that the FIM approach is only meaningful for sources with SNR $\gg 1$, so arbitarily weak sources are not permitted in the analysis.

Results

The parameter space for N slowly evolving binaries is $7N$ dimensional. This large dimensionality makes it difficult to carry out an exhaustive exploration of all the circumstances that can affect parameter estimation. These studies focus on sources that were close in frequency, and the other source parameters were chosen randomly. The frequency chosen for the first binary fixes the base frequency $f_{\rm base}$ of a data run. The frequencies of the remaining $N-1$ binaries were assigned frequencies of $f_{\rm base}+(i+x)f_m$, where $f_m = 1/\text{year}$ is the modulation frequency, i is an integer, and x is a random number between 0 and 1. Thus the remaining binaries are between i and $i+1$ modulation frequency units from the base frequency. Sky locations were chosen by two methods, the first being a random draw on $\cos(\theta) \in [-1, 1]$ and $\phi \in [0, 2\pi)$ with a fixed distance to the binary of 1 kiloparsec, the second being a random draw from a galactic distribution [67] of θ, ϕ, and binary distance. These two methods are referred to as all-sky and galactic draws, respectively. The values of the polarization ψ and orbital phase φ_o were randomly drawn between 0 and π and 0 and 2π respectively. The inclination ι was taken from a random draw on $\cos(\iota) \in [-1, 1]$, but with values outside of the range $\iota \in [1°, 179°]$ rejected to avoid the degeneracy in the gravitational wave produced by a circular binary viewed along its axis. Masses were taken to equal $0.5M_\odot$, modeling white dwarf binaries.

In what follows, we will typically quote our results in terms of uncertainty ratios. These ratios will compare the parameter resolution when one or more overlapping sources are present to the parameter resolution that would be possible if each source were isolated. To arrive at the absolute uncertainties one needs to multiply the uncertainty ratios by the isolated source uncertainties quoted in Chapter 3.

Varying Time of Observation

The current lifespan of LISA is estimated to be five years. The data show that the longer LISA observes the binary systems, the less the effect of confusion between sources. Figure 5.1 shows the median parameter uncertainty ratios (PURs) from all sky draws of two, three, and four binary systems located in a single modulation frequency bin based at 1 mHz as a function of the time of observation. Figure 5.2 shows the median geometric mean uncertainty ratio (GMUR) for the same data. Each set of PUR data in Figure 5.1 contains the uncertainty ratios for all seven parameters, and as can be seen in the figures, the uncertainty ratios for each of the seven parameters are closely related. This relationship holds true for other frequencies in the LISA band. It is this uniformity in the increase in parameter uncertainties that motivates the use of the GMUR as a global estimate for the individual PURs. Thus, subsequent results will be given in terms of the GMUR.

Figure 5.3 shows how the median GMUR increases with the number of binaries for one year of observation, at base frequencies of 0.1, 1, and 5 mHz. Figure 5.4 shows how extending the time of observation affects the GMUR. In both instances, the parameter uncertainties increase roughly exponentially with the number of overlapping sources.

Figure 5.5 shows the median of the magnitude of the correlation between the signals of two binary systems with a base frequency of 1 mHz taken from an all sky draw. As expected, the correlation magnitude falls off as $1/T_{obs}$ for observation times greater than a year (the numerator in (5.5) oscillates, while the denominator grows

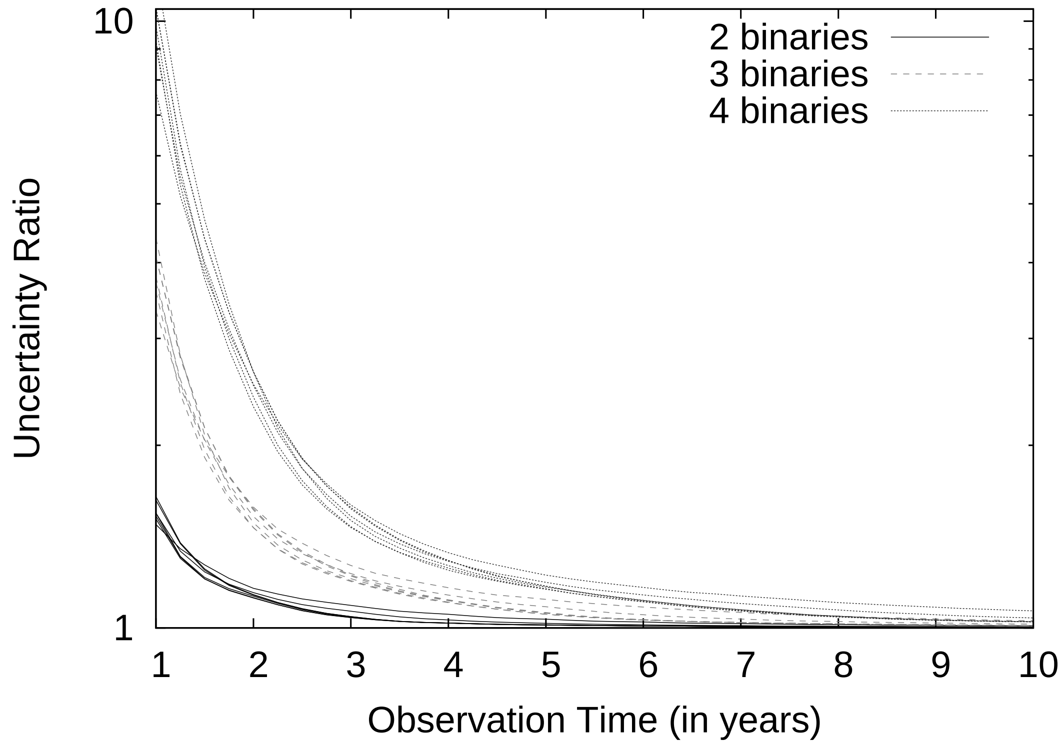

Figure 5.1: Median PUR versus time of observation for two, three, and four binaries with base frequencies of 1 mHz from an all sky draw.

as T_{obs}). The $1/T_{obs}$ fall-off in the signal correlation should be contrasted with the much faster fall-off in the GMUR seen in Figure 5.4. The reason for this difference in fall-off will become clear in the Uncertainty Ratios as a Function of Signal Correlation Section.

Uncertainty Ratios as a Function of Frequency Difference

Figure 5.6 shows how the median parameter uncertainties depend on the frequency separation when two binary systems are present. The base frequency, f_{base}, for the first binary was held fixed at 1 mHz, while the frequency of the second binary was randomly chosen between f_{base} and $f_{base} + if_m$ for i between 0 and 20. Figure 5.6 shows the plots for GMUR versus i for all sky and galactic draws for one year of

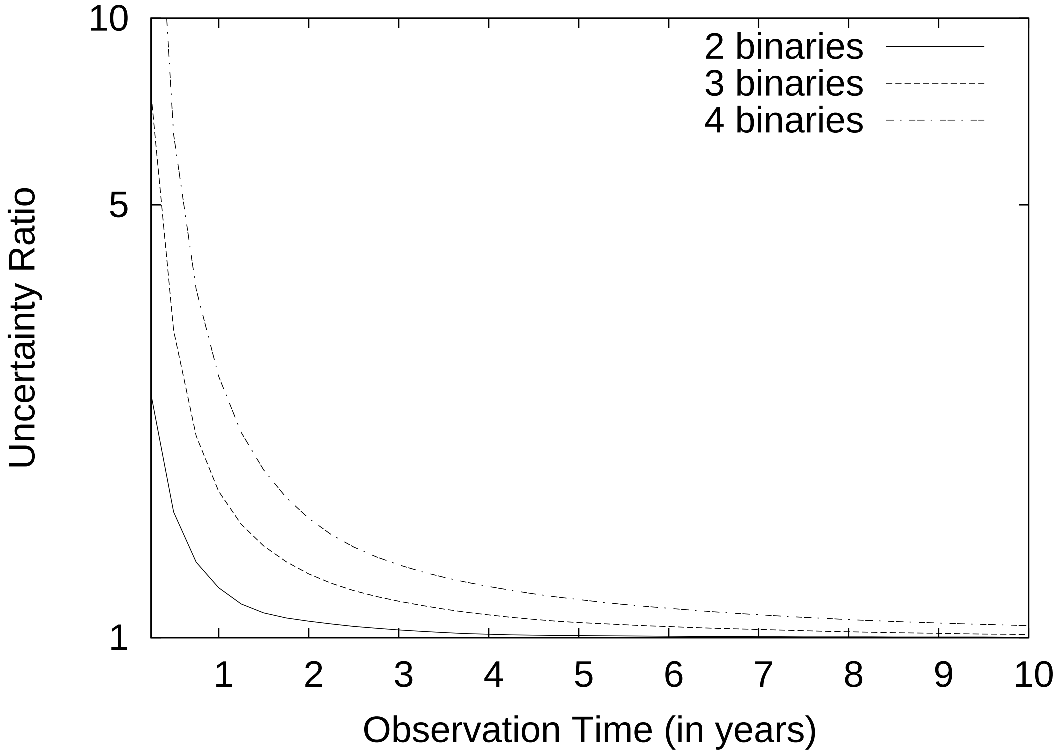

Figure 5.2: Median GMUR versus time of observation for two, three, and four binaries with base frequencies of 1 mHz from an all sky draw.

observation.

The uncertainty ratio drops rapidly as the frequency separation increases, approaching unity by the time the binaries are $\sim 5f_m$ apart. The frequency difference of $\sim 5f_m$ corresponds to the typical half-bandwidth of a source at 1 mHz. In other words, 1 mHz sources seperated by $> 5f_m$ have almost no overlap, and thus there is little source confusion for $i > 5$.

Uncertainty Ratios as a Function of Signal Correlation

Figure 5.7 plots the median GMUR versus the signal correlation for a galactic draw when two binaries are present within $\delta f = f_m$ of each other. The plot shows that the GMURs increase faster than exponentially as the signal correlation increases.

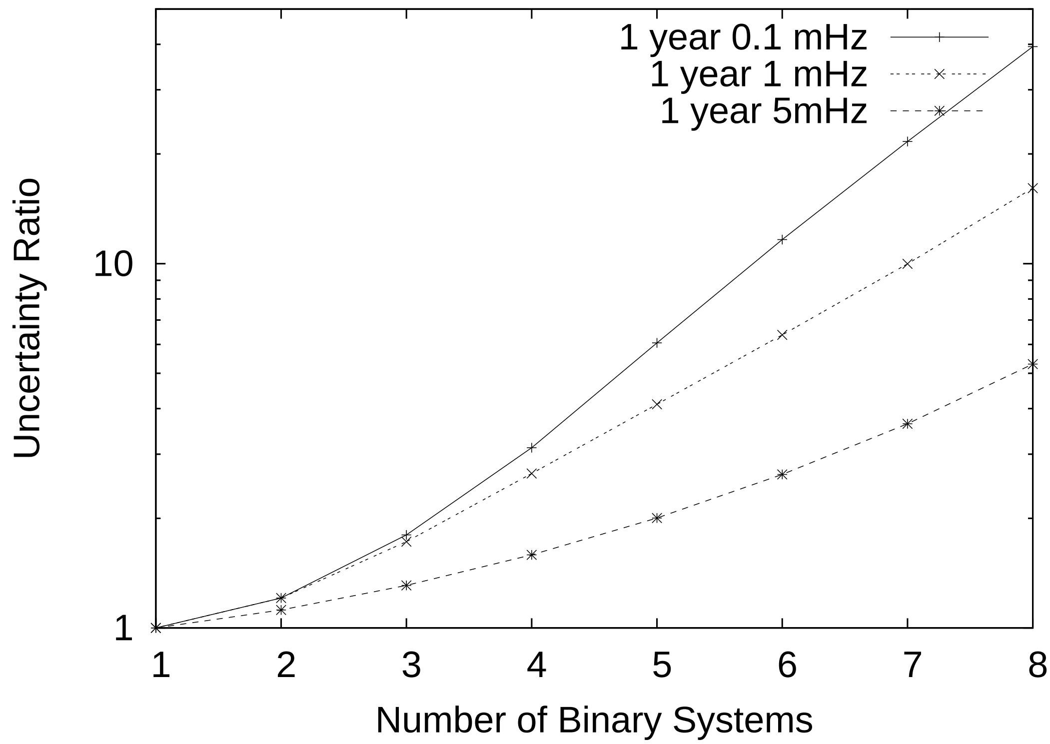

Figure 5.3: Median GMUR plotted against the number of binary systems for one year of observation at base frequencies of 0.1, 1, and 5 mHz.

While a link between signal correlation and uncertainty ratios is to be expected, the super-exponential nature of the relationship came as a suprise. Figures 5.8 and 5.9 show the spreads for correlation and GMUR in the same data run. As can be seen in the log-linear scale of Figure 5.9, the GMUR falls off exponentially from its most frequent value of 1.2.

While Figure 5.8 shows the results for two correlating binaries, its results can be extended to multiple correlating binaries. For example, in the figure one can see that 87% of the pairs of binaries have correlation magnitudes $|\kappa| < 0.5$. Thus with N binaries the probability that there will be a pair of binaries with $|\kappa| > 0.5$ grows as $1 - 0.87^{(N(N-1)/2)}$. As binaries overlapping in frequency are expected to number from tens to tens of thousands, the probability of having high correlation is very close to

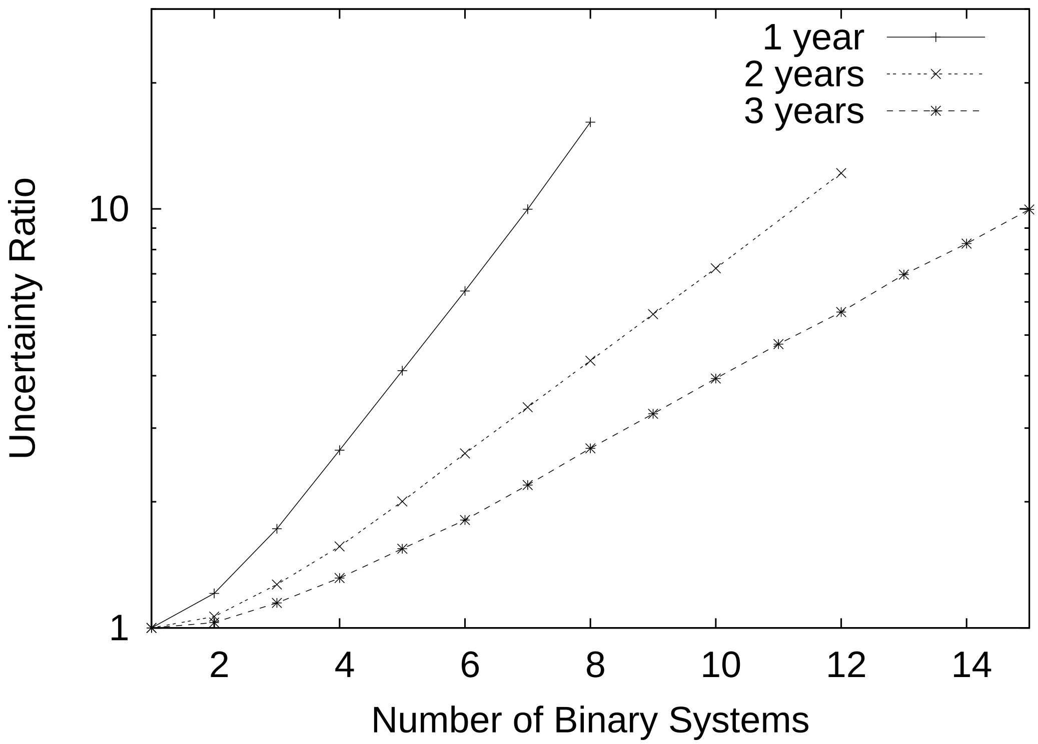

Figure 5.4: Median GMUR plotted against the number of binary systems for one, two, and three years of observation at a base frequency of 1 mHz.

1.

The rapid increase in the parameter uncertainties with signal correlation explains the results seen in the Varying time of observation Section. The $1/T_{obs}$ decrease in the signal correlation translates into a much faster decrease in the GMUR as a function of observation time.

Uncertainty Ratios as a Function of Base Frequency

Figure 5.10 shows how the parameter uncertainty ratios depend on base frequency. The plots are for all sky and galactic draws with two binaries for one year of observation. The base frequency, f_{base}, was varied between 0.01 and 10 mHz, while the frequency of the second binary was randomly chosen between f_{base} and $f_{base} + f_m$.

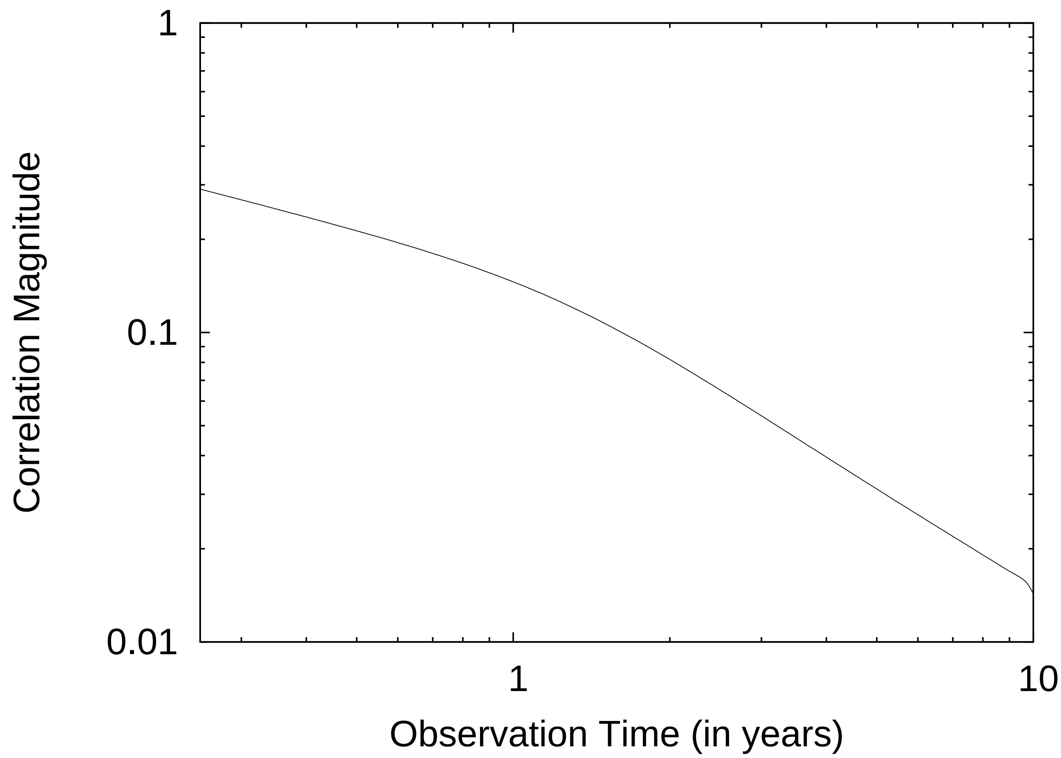

Figure 5.5: Median Correlation magnitude versus time of observation for two binaries with a base frequency of 1 mHz from an all sky draw.

The GMUR was found to be fairly constant below 1 mHz, followed by a rapid decrease at frequencies above 1 mHz. This behavior can be traced to the different Doppler shifts experienced by each source. The motion of LISA relative to a source located at (θ, ϕ) imparts a Doppler shift equal to

$$\delta f_D \simeq \pi \sin\theta \sin(2\pi f_m t - \phi) \left(\frac{f}{\text{mHz}}\right) f_m. \tag{5.10}$$

The magnitude of δf_D becomes comparable to the frequency resolution $\delta f = 1/T_{obs} =$ 1/year for $f \sim 0.3$ mHz. Sources that are well seperated in ecliptic azimuth ϕ will experience Doppler shifts that differ in sign, as LISA will be moving toward one source and away from the other. Thus, it is expected that the degree of source confusion

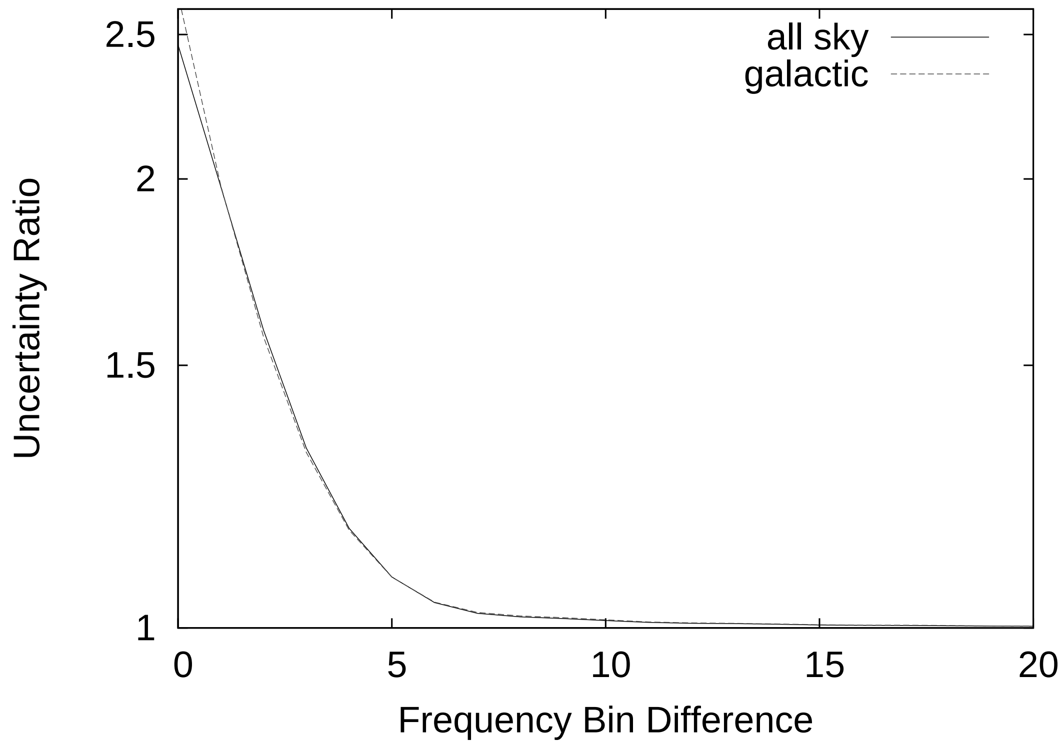

Figure 5.6: Median GMUR plotted against modulation frequency bin difference of two binaries from an all sky and a galactic draw with a base frequency of 1 mHz and one year of observation

will depend on the azimuthal separation of the two sources. These expectations are confirmed in Figure 5.11, where the dependence of the GMUR on azimuthal separation is plotted for base frequencies of 0.1 and 1 mHz. Unfortunately, most galactic sources are within $\sim 20^\circ$ degrees of each other in ecliptic azimuth, which helps explain the larger GMURs for the galactic distribution as compared to the all-sky distribution.

Information Theory Perspective

There have been several attempts to understand LISA source confusion in the framework of information theory [74, 75, 76]. According to Shannon [77], the maximum amount of information that can be transmitted over a noisy channel with

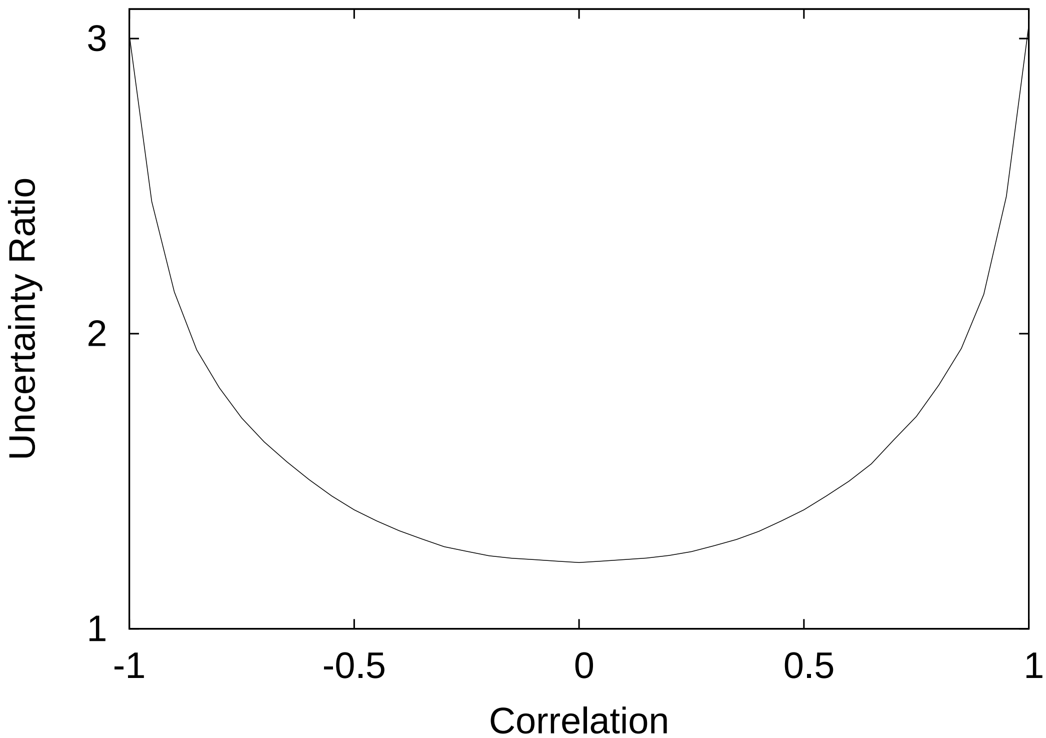

Figure 5.7: The median GMUR plotted against signal correlation of two binaries from a galactic sky draw with a base frequency of 1 mHz and one year of observation

bandwidth B in a time $T_{\rm obs}$ is

$$I_S = B\,T_{\rm obs}\log_2\left(1+\frac{P_h}{P_n}\right). \tag{5.11}$$

Here P_h and P_n are, respectively, the signal and noise power across the bandwidth. (Note: The SNR ratio is related to these quantites by $\text{SNR}^2 \simeq B\,T_{\rm obs}P_h/P_n$). As an example, a typical bright galactic source at 1 mHz with $P_h/P_n \sim 10$ can transmit at most $I_S \simeq 70$ bits of information in one year via the two independent LISA data channels S_I and S_{II}. In practice the actual information content will be somewhat less due to sub-optimal encoding. It takes $\log_2(x/\Delta x)$ bits of information to store a number x to accuracy Δx, thus the total amount of information required to describe

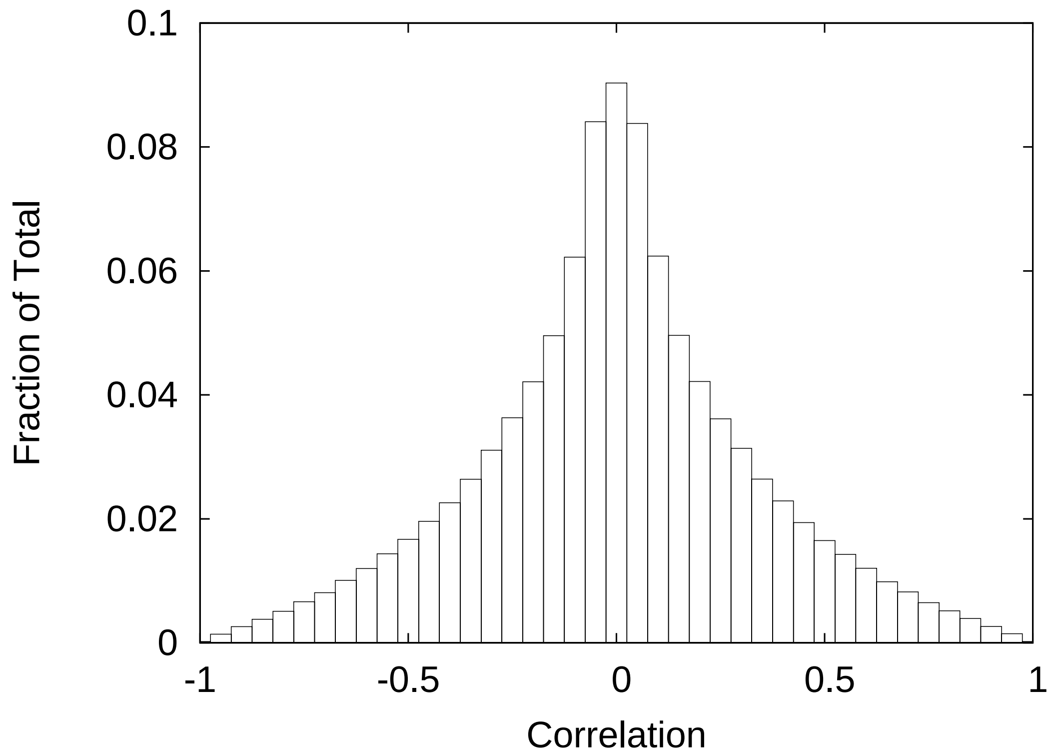

Figure 5.8: A histogram of the correlation of two binaries from a galactic draw with a base frequency of 1 mHz and one year of observation

a monochromatic binary to a precision $\Delta\vec{\lambda}$ is

$$\begin{aligned} I_\Delta = \log_2\left(\frac{f}{\Delta f}\right) + \log_2\left(\frac{A}{\Delta A}\right) + \log_2\left(\frac{\pi}{\Delta\theta}\right) \\ + \log_2\left(\frac{2\pi}{\Delta\phi}\right) + \log_2\left(\frac{\pi}{\Delta\psi}\right) \\ + \log_2\left(\frac{\pi}{\Delta\iota}\right) + \log_2\left(\frac{2\pi}{\Delta\gamma}\right) \end{aligned} \tag{5.12}$$

Using the results in Figure 3.1 for a bright galactic source at 1 mHz with $P_h/P_n \sim 10$, it is found that $I_\Delta \simeq 45$, which indicates that the encoding, while sub-optimal, is quiet respectable.

One might hope that information theory could be used to predict some of the other results described in Section 5. Unfortunately, this proves not to be the case,

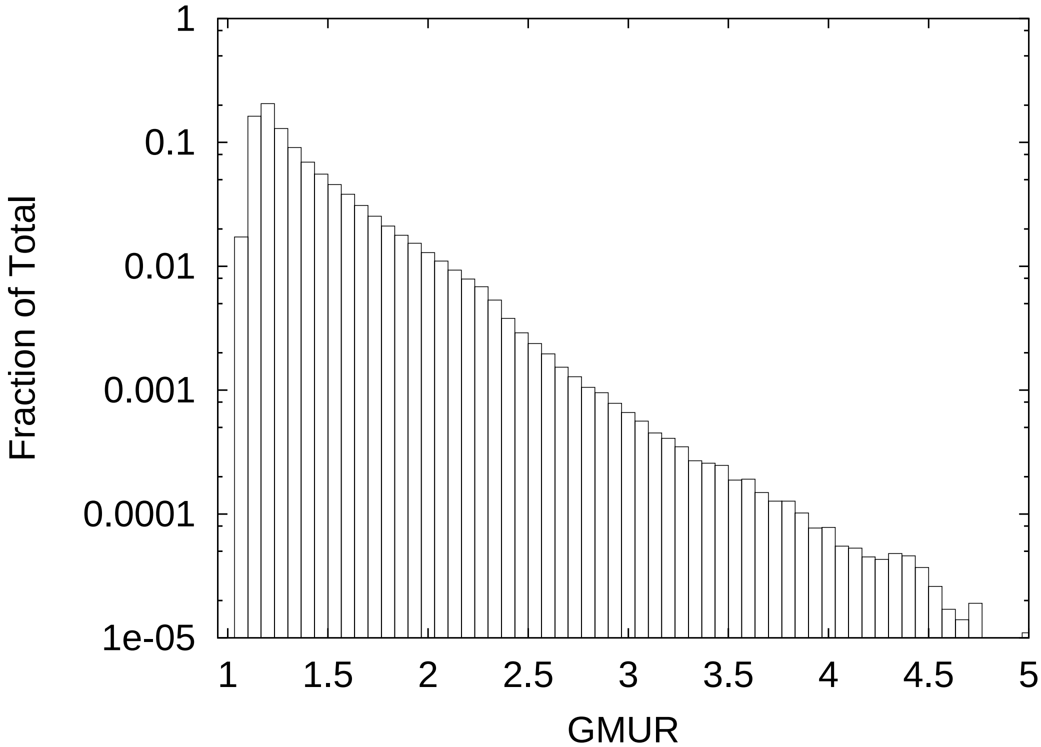

Figure 5.9: A histogram of the GMUR of two binaries from a galactic draw with a base frequency of 1 mHz and one year of observation

for while information theory sets limits on how well one might do, it sets no limits on how poorly. The scaling of the GMUR as a function of the number of sources is a good example. The information needed to localize N galactic sources is equal to

$$I_\Delta(N) = \log_2\left(\frac{V_{7N}}{\Delta V_{7N}}\right) \tag{5.13}$$

where ΔV_{7N} is given by (5.4) and $V_{7N} = V_7^N$ is the volume of the $7N$ dimensional parameter space. Requiring that $I_\Delta(N) < I_S(N)$ yields a lower bound on the geometric mean of the parameter uncertainties (GMU) of

$$\text{GMU} > \frac{V_7^7}{(1 + P_s/P_n)^{BT/7N}}. \tag{5.14}$$

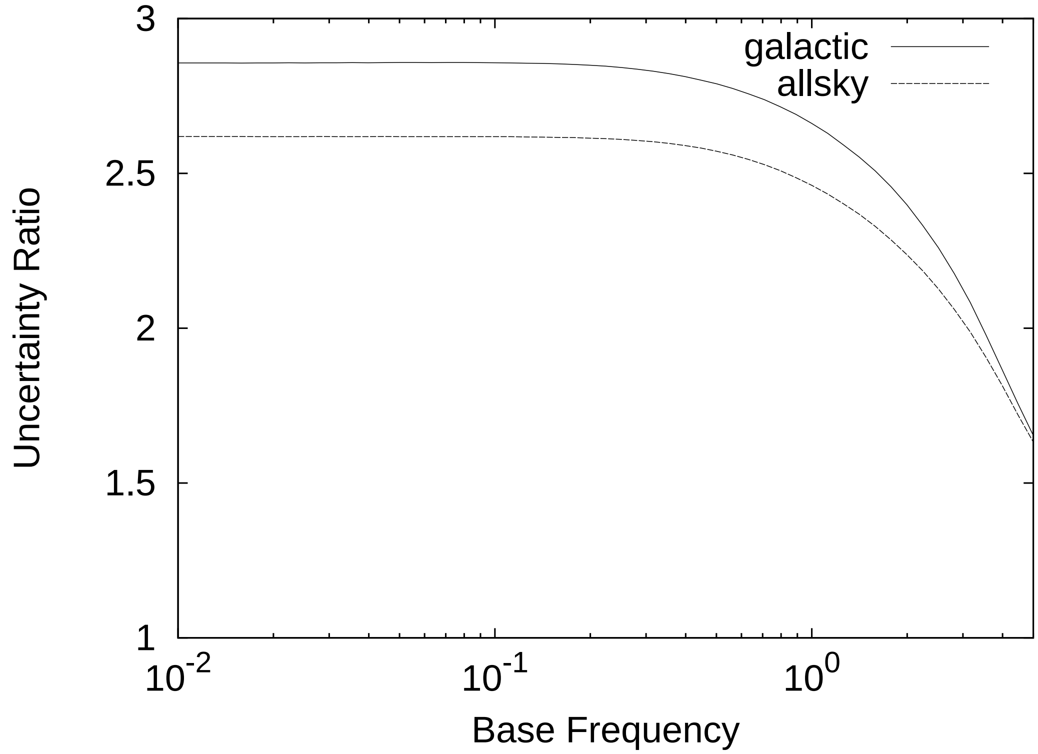

Figure 5.10: Median GMUR plotted against base frequency of two binaries from all sky and galactic draws with one year of observation

Consider first the case of N isolated binaries of similar brightness. The power ratio P_s/P_n will be independent of N, and the bandwidth will be the sum of the bandwidths of each signal, so B will grow linearly with N. Thus, Equation (5.14) implies that for isolated binaries, the parameter uncertainties will be independent of the number of sources. This is indeed the case, as the FIM is block diagonal for isolated sources. Now consider the case of N overlapping binaries sharing the same fixed bandwidth B. The power P_s will grow approximately linearly with N, leading to the prediction that the GMU must grow faster than $(1/N)^{(1/N)}$. The actual scaling seen in Figures 5.3 and 5.4 tell a different story, with the GMUR increasing as $\sim N^N$. Thus, while information theory provides a lower bound on how quickly the parameter uncertainties must increase with the number of overlapping sources, the bound is too

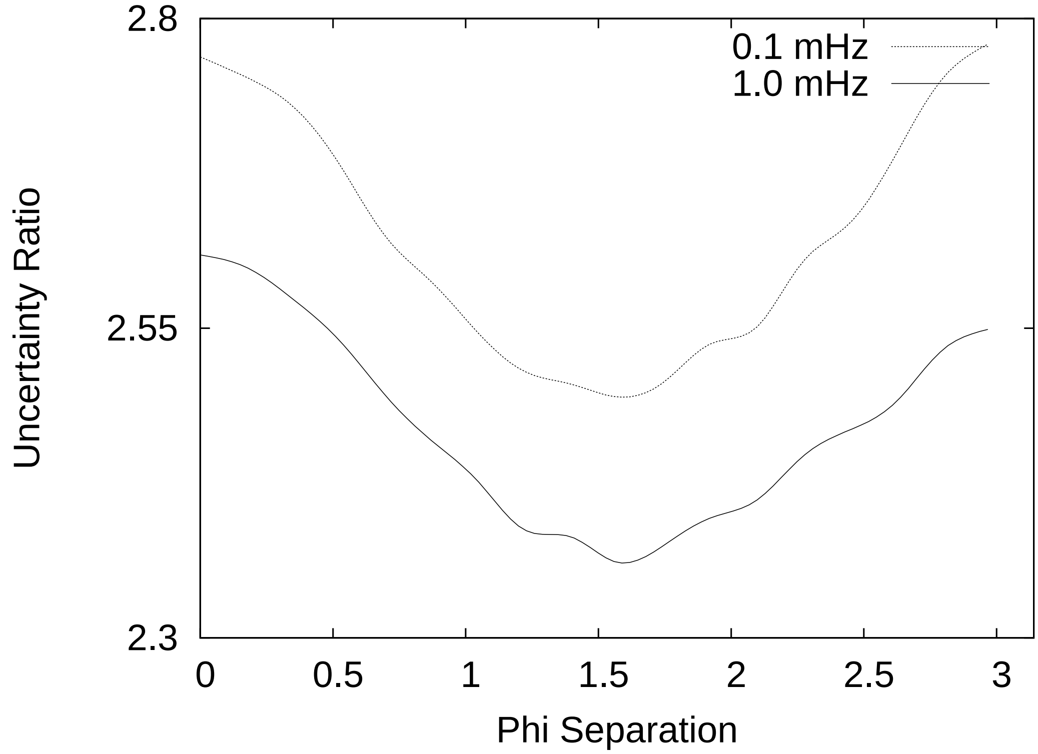

Figure 5.11: Median GMUR plotted against azimuthal sky separation for two pairs of binaries from a galactic draw with one year of observation with base frequencies of 0.1 and 1 mHz.

weak to be of any real use. The weakness of the bound is probably related to how poorly the information for multiple overlapping binaries is encoded.

CHAPTER 6

GENETIC ALGORITHMS AS A SEARCH METHOD

Introduction

This chapter focuses on the use of GAs as a method of searching the LISA data stream to determine the source parameters of the monochromatic binary systems used to create the simulated LISA signal. The bulk of this work was first published in [56].

Previous approaches to the extraction of parameters from the LISA data stream have used several methods. Grid based template searches using optimal filtering provide a systematic method to search through all possible combinations of gravitational wave sources, but the computational cost of such a search appears to make it unfeasible [78]. Other techniques applied to simulated LISA data involve iterative refinement of a sequential search of sources [49, 51], a tomographic approach [79], global iterative refinement, and ergodic exploration of the parameter space such as MCMC methods [54]. At this time, however, it is not clear which of these techniques, or which combination of techniques will provide the best solution to the so-called Cocktail Party Problem discussed in Chapter 3.

Here an application of the method of genetic algorithms [55] to the challenge of extracting parameters from a simulated LISA data stream containing multiple monochromatic gravitational wave sources is presented. The strength of this method lies in its searching capabilities. Several studies [48, 60, 80, 81, 82, 83, 84, 85] have indicated that the confusion noise may dominate instrument noise at the low end of the LISA frequency range, so that other sources of interest may be buried beneath the confusion background. For this reason a key goal of LISA data analysis is to reduce the level of the confusion noise as much as possible. GAs might be used as the first step in dealing with the many thousands of these low frequency, effectively monochromatic sources that constitute confusion background. The initial solution could then be handed off to a MCMC algorithm similar to the one discussed in Chapter 7, which specializes in determining the nature of the posterior distribution function.

The chapter is organized as follows: After the executive summary, basic terms are defined, a bare-bones algorithm is introduced, and succeeding layers of complexity are added to this algorithm. An emphasis is placed on developing an efficient algorithm which is robust enough to handle the entire low frequency regime of the LISA detector. Applications of the advanced algorithms to multiple source cases are shown, followed by an GA incorporating directed search methods. The chapter concludes with a discussion of future improvements and plans for the application of genetic algorithms to LISA data analysis.

Chapter Executive Summary

This chapter is an introduction to and an application of a genetic algorithm to the search of gravitational wave source parameters. It will show that the method is a feasible search method capable of handling multiple sources in a restricted frequency range. While the bare bones algorithm will demonstrate the ability to function as a search method, advancements in the algorithm will provide a large increase in computational efficiency. The advanced genetic-genetic version of algorithm is seen to perform solidly even in the low SNR regime.

Genetic Search Algorithms

The fundamental idea behind a genetic algorithm is the survival of the fittest.

It is because of this that genetic algorithms are often referred to as evolutionary algorithms, though Darwin [86] would probably have considered GAs as "Variation under Domestication" since the breeding is directed toward a predetermined goal. Through the process of continually evolving solutions to the given problem, genetic algorithms provide a means to search the large parameter space that exists in the low frequency region of the LISA band.

A few definitions are in order before delving into the applications of genetic algorithms to LISA data analysis. These definitions will refer to a hypothetical search of the LISA data stream for N monochromatic gravitational wave sources. The search will take advantage of the F-statistic to reduce the search space to $3N$ parameters. The hypothetical search will also involve the use of n simultaneous, competing solution sets.

An **organism** is a particular $3N$ parameter set that is a possible solution for the source parameters.

A **gene** is an individual parameter within an organism.

A **generation** is the set of all n concurrent organisms.

Breeding or cross-over is the process through which a new organism is formed from one or more organisms of the previous generation.

Mutation is a process which allows for variation of a organism as it is bred from the organisms of the previous generation.

Elitism is the technique of carrying over one or more of the best organisms in one generation to the next generation.

A simplified genetic algorithm begins with a set of n organisms that comprise the first generation. The genes of this generation may be chosen at random or selected through some other process. The organisms of each generation are checked for fitness, and those with the best fitness are more likely to breed, with mutation, to form the organisms of the next generation. With passing generations the organisms tend toward better solutions to the source parameters. The F-statistic is used to measure the fitness of each organism.

Basic Implementation

For these investigations source frequencies were chosen to lie within the range $f \in [0.999995, 1.003164]$ mHz. This range spans 100 frequency bins of width $\Delta f =$ 1/year. Amplitudes were restricted to the range $A \in [10^{-23}, 10^{-21}]$. By use of the F-statistic the searches are reduced to frequency f, and sky location θ and ϕ. For a detailed description of the F-statistic and its use in reducing the search space see Chapter 3.

A simple approach is to represent the values of each search parameter with binary strings. The length of the strings determines the precision of the search, e.g. representing θ with a binary string of 8 digits gives precision to 0.7°. Resolution is given by, (parameter range)$/2^L$, where L is the length of the binary string. Such a binary representation allows for ease of mutation and breeding. This work employed binary strings of length $L = 16$ for f, $L = 13$ for θ and $L = 14$ for ϕ.

In this basic scheme, the parent's parameter strings are first mutated, and then the mutated gametes are bred together. Simple mutation consists of flipping the binary digits of the parent's parameter strings with probability PMR, the parameter mutation rate. A large PMR will tend to result in more variation in the gametes, and thus the offspring, while a small PMR will lessen variation, resulting in more offspring that resemble their parents.

This study uses a breeding pattern known as 1-point crossover, which consists of the combination of complimentary sections of the binary strings of two parent organisms. The cross-over point can be chosen at random or fixed in advance. For simplicity, a fixed cross-over with the cross-over point occurring at the midpoint of the strings. As an example, Table 6.1 shows the breeding of a parameter represented by strings that are 8 digits long.

Table 6.1: Midpoint crossover for an 8 bit string

Parent 1	0100	1110
Parent 2	0011	0011
Offspring	0011	1110

The basic search will use 10 organisms in each generation. The first generation has the genes of its organisms chosen at random from their respective ranges. The probability of each of these organisms being chosen for reproduction is proportional to its likelihood, $\mathcal{L}$ (known as fitness proportionate breeding). Mutated gametes are formed using a PMR of 0.04, and are bred using a single midpoint crossover.

Figure 6.1 shows trace plots of the log likelihood, frequency, θ, and ϕ for a source with SNR = 15.4464 and parameters: $A = 1.97703 \times 10^{-22}$, $f = 1.000848032$ mHz, $\theta = 1.2713$, $\phi = 5.34003$, $\iota = 2.73836$, $\psi = 1.43093$, and $\gamma_o = 5.59719$ (it is this source that will be used repeatedly throughout the paper). The plotted values were for the organism with the best fit in each generation. As can be seen the parameters are well determined with even this basic scheme, though the noise in the data stream pushes them off their true values. The parameter values are shifted by $\delta f = -1.5 \times 10^{-9}$ Hz, $\delta\theta = 2.9°$ and $\delta\phi = -1.5°$ from their input values. These shifts are consistent with the error predictions from a Fisher Information Matrix analysis: $\Delta f = 1.7 \times 10^{-9}$ Hz, $\Delta\theta = 3.5°$ and $\Delta\phi = 1.9°$. The cost of the search is measured in terms of the number of calls to the F-statistic routine and is given by $\$ = n \times g$, where g is the generation number. Typical runs of the basic genetic algorithm cost $\$ = 32650$ calls. This should be compared to a grid based search across the same frequency range, which, for a minimal match of MM = 0.9, would require $\$ = 110,000$ calls to the F-statistic routine (this value is $2^{3/2}$ larger than that quoted in Ref. [54] as the earlier calculations used a noise level that was $\sqrt{2}$ larger than the LISA baseline due to a mix up between one and two sided noise spectral densities).

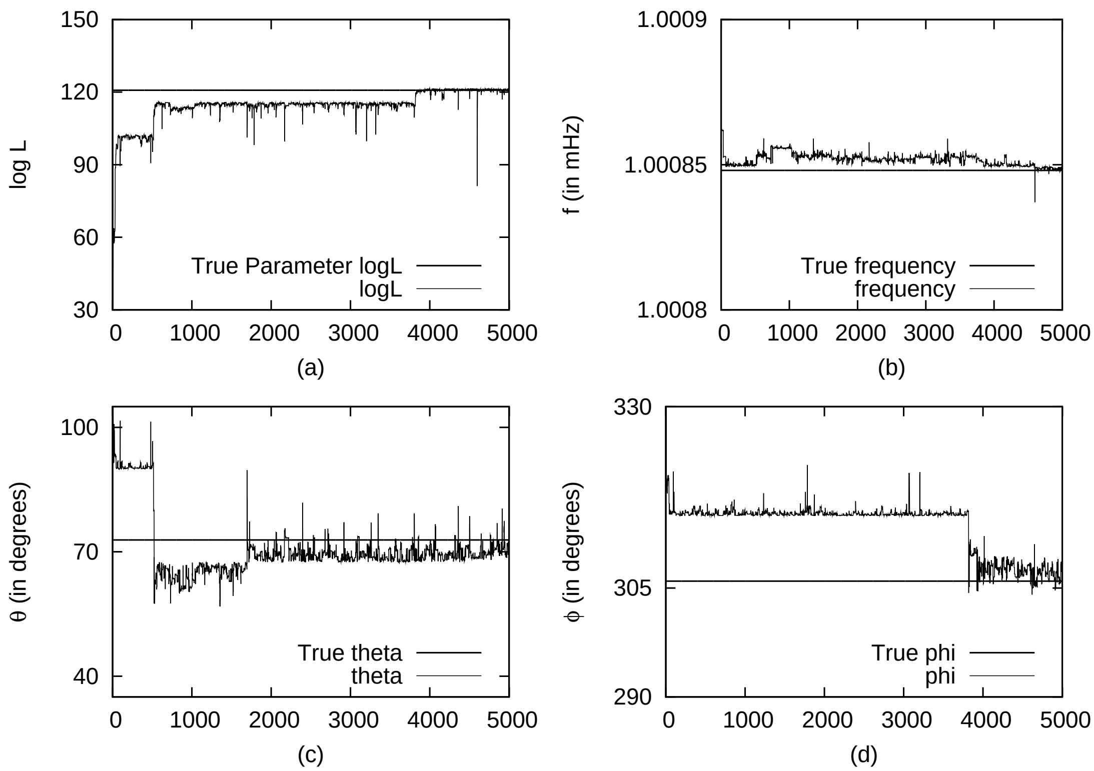

Figure 6.1: Basic Algorithm: Trace plots for (a) log likelihood, (b) frequency, (c), θ, and (d) ϕ for the basic implementation of a genetic algorithm. The y-axes are the parameter values and log likelihoods of the best fit organism for each generation. The x-axes are generation number.

While the basic algorithm is sufficient for finding a solution, it is not efficient. Next adjustments to the algorithm that will improve its efficiency are introduced and analyzed, making the algorithm considerably cheaper than a grid-based search.

Aspects of Mutation

In the previous example the PMR was set at the fairly low value of 0.04. Figure 6.2 shows trace plots for the same search, but with PMR = 0.1. While the PMR = 0.04 example shows a tendency for small deviations from the improving solutions, the larger PMR search allows large swings in the solution away from a good fit to the true source parameters. On the other hand, Figure 6.3 shows how a small PMR (0.001) can cause the rate of progress to be greatly slowed. A small mutation rate slows the exploration of the likelihood surface.

As these examples show, choosing the proper PMR can have a significant effect on the efficiency of the algorithm. Knowing which value is the proper choice a priori is impossible. Furthermore, at different phases of the search, different values of the PMR will be more efficient than those same values at other phases. Early on in the search a large PMR is desirable for increased exploration. Once convergence to the solution has begun, a smaller PMR is preferable, to prevent suddenly mutating away from the solution. One can imagine a process which changes the PMR in a manner analogous to the simulated annealing process, where at the start the PMR high (hot) and it is lowered (cooled) it in succeeding generations. In fact, this process in sometimes called simulated annealing in the GA literature. Figure 6.4 shows trace plots for the same source, using a genetic (PMR) simulated annealing scheme given by:

$$\mathrm{PMR} = \begin{cases} \mathrm{PMR_f} \left(\frac{PMR_\mathrm{i}}{PMR_\mathrm{f}}\right)^{\frac{g_\mathrm{cool}-g}{g_\mathrm{cool}}} & 0 < g < g_\mathrm{cool} \\ \mathrm{PMR_f} & g \geq g_\mathrm{cool} \end{cases} \tag{6.1}$$

where $\mathrm{PMR_i} = 0.2$, $\mathrm{PMR_f} = 0.01$, g is the generation number, and $g_\mathrm{cool} = 1000$ is the last generation of the cooling process. The best choice of values for this scheme is again impossible to know a priori. In the Giving More Control to the Algorithm Section one will see how "Genetic Genetic Algorithms" are able to provide a natural solution to this problem.

The Effect of Organism Number on Efficiency

While choosing the PMR is one degree of freedom in the basic schema, another

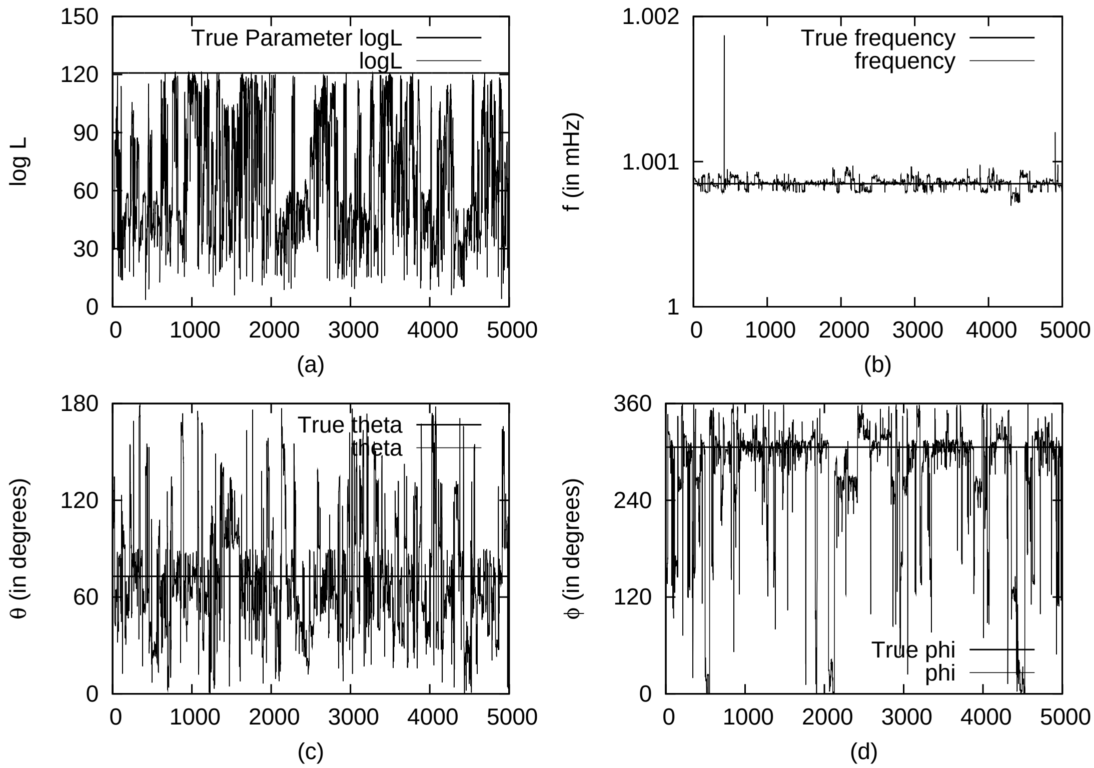

Figure 6.2: Large Mutation Rate: Trace plots for (a) log likelihood, (b) frequency, (c), θ, and (d) ϕ for the basic implementation of a genetic algorithm with PMR $= 0.1$. The y-axes are the parameter values and log likelihoods of the best fit organism for each generation. The x-axes are generation number.

is the number of organisms used in the search. One can look at how the choice of the number of organisms effects the efficiency of the algorithm. The efficiency is inversely related to the computational cost \$, which is measured by the number of calls to the function calculating the F-statistic (where the bulk of calculations for an organism are performed), which occurs once per newly formed organism. For example, in Figure 6.1 there are 10 organisms in the search and the search surpasses the true parameter log likelihood value at 3851 generations. Thus its computational cost is $\$ = 38510$ (function calls).

The data in Figure 6.5 shows the interplay of the number of organisms with the PMR (held constant within each data run) and their effects on the computational

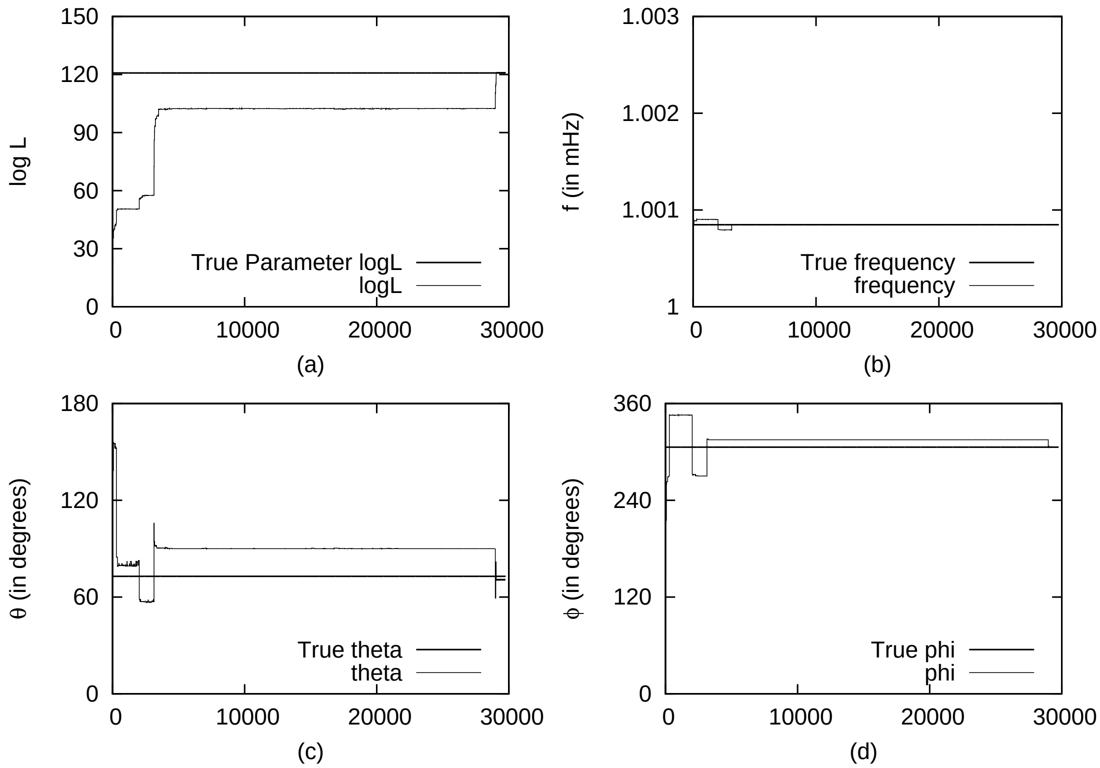

Figure 6.3: Small Mutation Rate: Trace plots for (a) log likelihood, (b) frequency, (c), θ, and (d) ϕ for the basic implementation of a genetic algorithm with PMR = 0.001. The y-axes are the parameter values and log likelihoods of the best fit organism for each generation. The x-axes are generation number.

cost. One would expect that relatively large PMRs would be less efficient as was seen in the Aspects of Mutation Section (and will show up in Figure 6.7). The size of the effect, however, is modified by the number of organisms in the search. For example, one can find from Figure 6.5 that the minimum cost ($\$ = 4492$) for a 20 organism search occurs when PMR = 0.1, however for 400 organisms in the search the minimum cost ($\$ = 7490$) is at PMR = 0.14.

The addition of more organisms in the search provides a kind of stability to the system that decreases the chances of mutating away from good solutions. With just a handful of organisms, and a large PMR, the chances are higher of each organism undergoing a large mutation in at least one parameter. However, with hundreds of

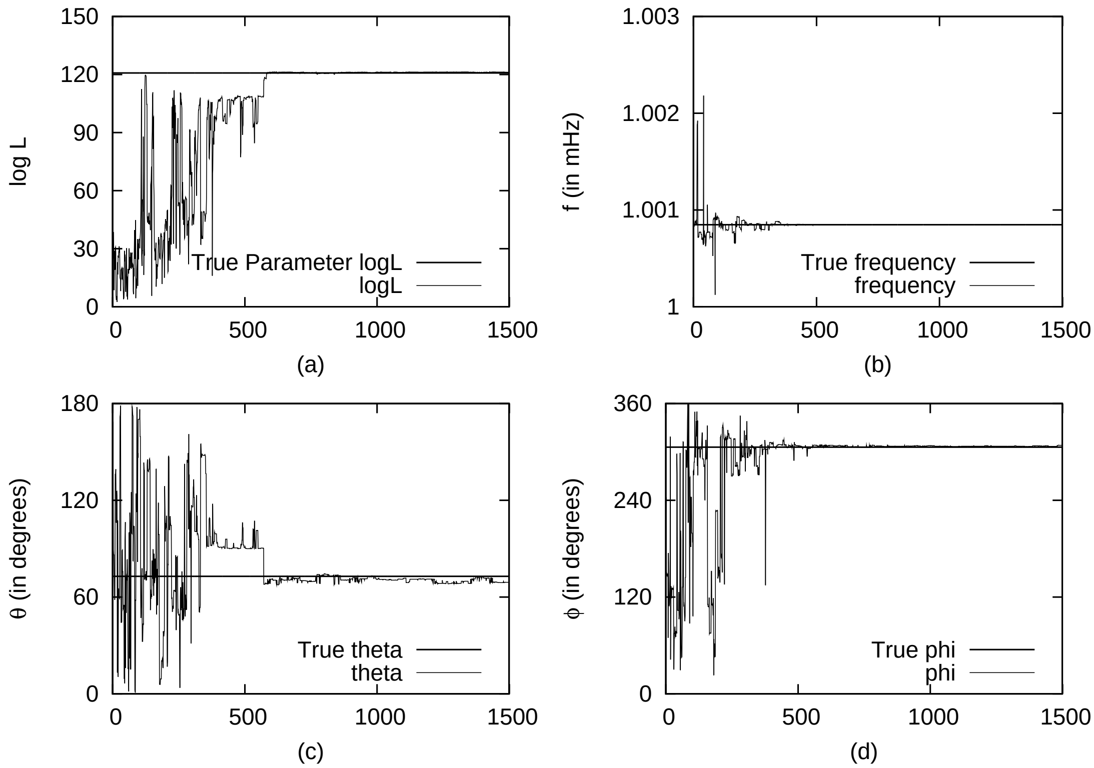

Figure 6.4: Genetic Simulated Annealing: Trace plots for (a) log likelihood, (b) frequency, (c), θ, and (d) ϕ for the basic implementation of a genetic algorithm with the inclusion of genetic simulated annealing. The y-axes are the parameter values and log likelihoods of the best fit organism for each generation. The x-axes are generation number.

organisms the probability of all organisms undergoing such a mutation drops appreciably. Then in the succeeding generation, those organisms that remained a good fit are much more likely to breed the offspring of the next generation. However, this does not hinder great leaps forward. To illustrate this point the data shown in Figure 6.1 is used. In going from the 7^{th} to the 8^{th} generation the value of the likelihood of the best fit organism jumps from 1.48×10^{13} to 6.02×10^{20}. As the probability of breeding is set by the value of the organism likelihood, that new best fit organism is going to be the primary breeder of the next generation (though it is possible that a second organism has also jumped to a point in parameter space with a similar likelihood value).

Increasing the number of organisms not only provides this stabilizing effect, it also provides more chances per generation for improvements due to mutations. One cannot, however, simply throw more organisms at the problem without paying a price; that price will be an eventual drop in efficiency. As an extreme example, imagine using the basic scheme described in the Basic Implementation Section and putting 40000 organisms into the search. Even if one of the randomly chosen organisms matched the best fit parameters, the computational cost ($\$ = 40000$) is already larger than the cost of using 10 organisms ($\$ = 38510$). Figure 6.5 provides a snapshot of the how this choice effects efficiency.

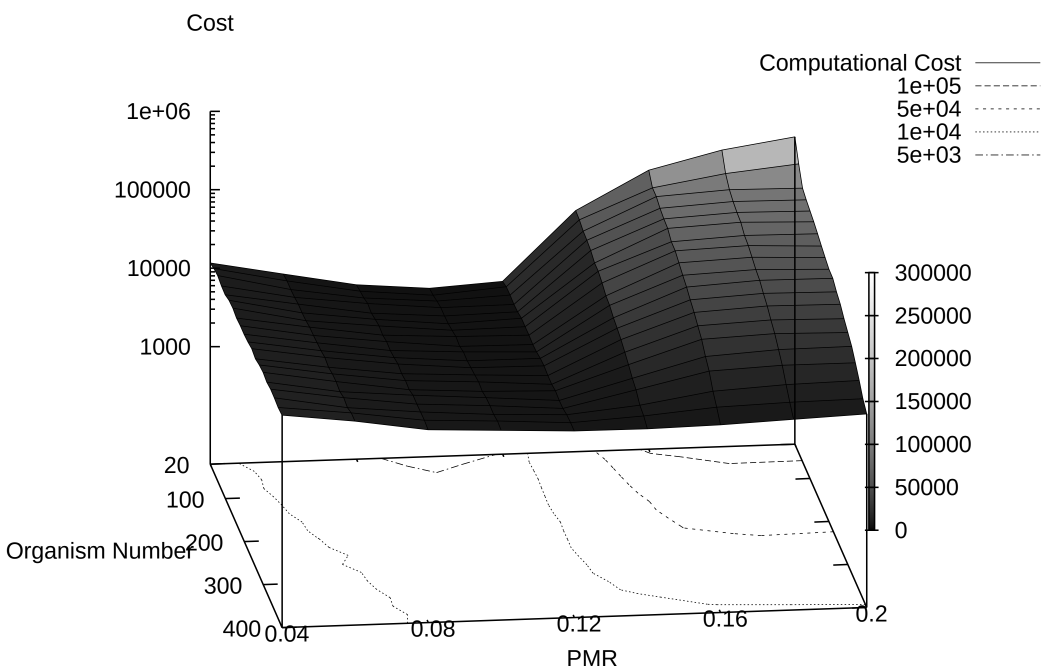

Figure 6.5: Average Computational Cost as a function of PMR and the number of organisms. The z-axes is the average computational cost calculated from 1000 searches.

Elitism

Elitism is akin to cloning. It allows for a perfect copy of an organism or organisms to be carried over into the next generation. Including elitism is another way to provide a stabilizing force across generations. This allows for a larger PMR to enhance exploration without the danger of moving off the best fit solution.

Figure 6.6 shows trace plots for the nominal source with PMR = 0.1 and a single elite organism being cloned at each generation. As expected there is increased exploration (compared to results shown in Figure 6.1) due to the larger PMR, but unlike the results shown in Figure 6.2, convergence is now helped by the cloned organism.

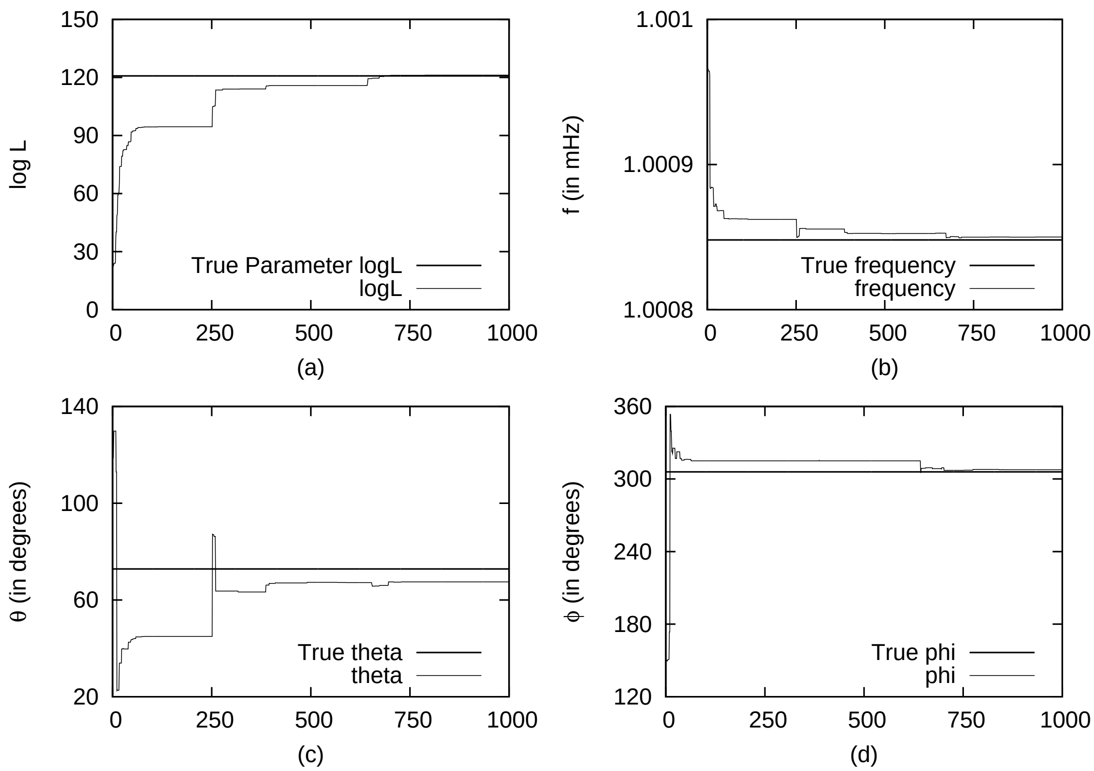

Figure 6.6: Elitism: Trace plots for (a) log likelihood, (b) frequency, (c), θ, and (d) ϕ for the basic implementation of a genetic algorithm with PMR = 0.1 and single organism elitism. The y-axes are the parameter values and log likelihoods of the best fit organism for each generation. The x-axes are generation number.

Figure 6.7 shows a plot relating the average computational cost to the PMR for the case of no elitism, and the case where a single organism is cloned. Computational cost is now derived from the average number of newly formed organisms (note: a cloned organism does not increase computational cost, as all of its associated values are already known). The plot shows the average computational cost of 100 searches, using 20 organisms, of a given source (SNR = 19.2335 and parameters: $A = 1.61486e-22$, $f = 1.003$ mHz, $\theta = 0.8$, $\phi = 2.14$, $\iota = 0.93245$, $\psi = 2.24587$, and $\gamma_o = 5.29165$). As was expected, elitism has allowed for a larger PMR, compared to the zero elitism case, increasing the parameter space exploration without sacrificing efficiency.

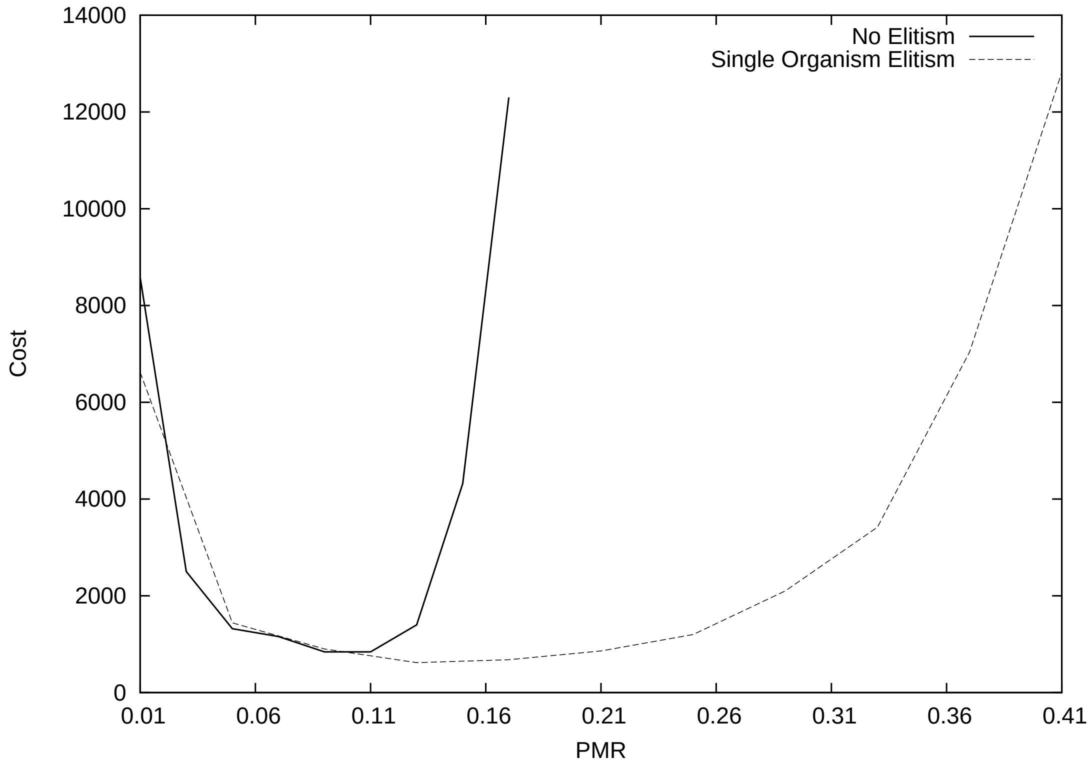

Figure 6.7: Average Computational Cost for no elitism and single organism elitism. Data points are determined by the average of 100 distinct searches.

If one decides to use elitism there is the additional choice of how many elite organisms will be cloned at each generation. At one extreme all organisms are cloned, in which case there is no exploration beyond the first generation. At the other extreme of no elitism the algorithm is unstable against large PMR values, as was seen in Figure 6.2. There is a balance to be struck between the amount of elitism and the size of the PMR that will provide the most efficient scheme, but the exact nature of the balance can depend on the nature of the search. A solution to this problem is described in the Giving More Control to the Algorithm Section.

Simulated Annealing

Simulated annealing is a technique that effectively makes the detector more noisy, thus lessing the range of the likelihood function. This increases the probability of choosing poorer sources for reproduction, which allows for a more thorough exploration of the likelihood surface. Think of the likelihood as a partition function $Z = C\exp(-\beta E)$, in which the role of the energy is played by the log likelihood, $E = (s - h|s - h)$, and β plays the role of the inverse temperature. Heating up the system (lowering β) lowers the likelihood range, providing for increased exploration. Starting hot, a power law cooling schedule is used.

$$\beta = \begin{cases} \beta_0 \left(\frac{1}{2\beta_0}\right)^{g/g_{\mathrm{cool}}} & 0 < g < g_{\mathrm{cool}} \\ \frac{1}{2} & g \geq g_{\mathrm{cool}} \end{cases} \tag{6.2}$$

where β_0 is the initial value of the inverse temperature, g is the generation number, and g_{cool} is the last generation of the cooling process (subsequent generations have $\beta = 1/2$). As the likelihood is a sharply peaked function, it was found that for a single source an initial value of $\beta_0 \sim 1/100$ was sufficient to speed the process. For multiple source searches increasing that by factors of 3 to 5 produced more efficient explorations. Similarly, for multiple sources an increase in g_{cool} was needed to properly explore the surface. This increase scaled roughly linearly with the number of sources.

This mode of simulated annealing, which will be referred to as standard simulated annealing, is markedly different than the genetic version of simulated annealing discussed in the Aspects of Mutation Section. Standard simulated annealing alters the search space, using the heat/energy to smooth the likelihood surface, whereas in genetic simulated annealing the search space was left unchanged and the heat/energy of the organisms was increased via the larger PMRs.

Figure 6.8 shows trace plots of the log likelihood, frequency, θ, and ϕ searching for the same source as in Figure 6.4. The only change between the two examples is the type of annealing process. For this run PMR $= 0.04$, $\beta_0 = 1/100$, and $g_{\text{cool}} = 300$.

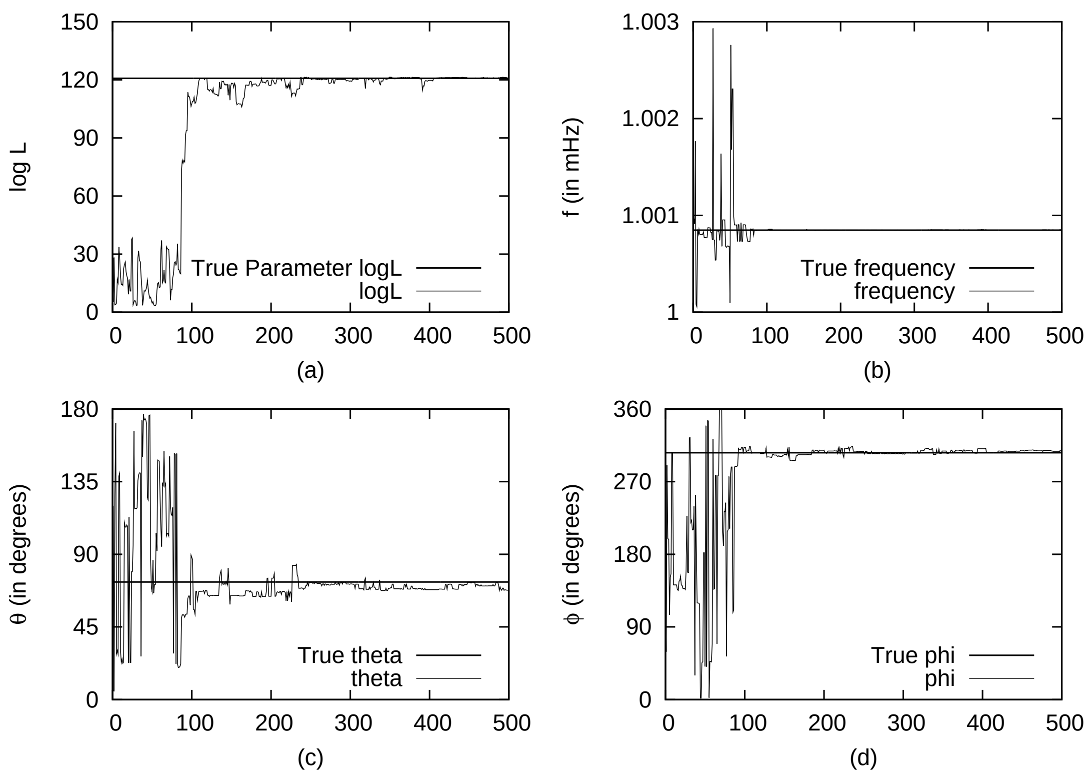

Figure 6.8: Standard Simulated Annealing: Trace plots for (a) log likelihood, (b) frequency, (c), θ, and (d) ϕ for the basic implementation of a genetic algorithm with the inclusion of standard simulated annealing and PMR $= 0.04$. The y-axes are the parameter values and log likelihoods of the best fit organism for each generation. The x-axes are generation number.

Giving More Control to the Algorithm

In the previous examples, choices were required as to what PMR or which degree of elitism should be used with a particular source to provide the most efficient search. In making those choices, one is searching for a solution that depends on the information in the data stream. Just as the power of the genetic algorithm can be used to search for the parameters of the gravitational wave sources that contribute to the data stream, one can also use that same power to search for efficient values for PMR or elitism.

Treating the PMR, elitism, or other factors in the genetic algorithm like a source parameter these factors can be elevated, or one might say demoted, to the same level as the source parameters. This was mentioned at then end of the Aspects of Mutation Section, and here it is implemented for the PMR. The initial PMR for each organism is chosen randomly, and the PMR for each organism in the next generation is bred just as f, θ, and ϕ are, based on organism fitness. This changes the nature of the algorithm from a simple genetic algorithm to a genetic-genetic algorithm (GGA), in which a factor, or factors, determining the search for the source parameters evolve along with the organisms.

Figure 6.9 shows trace plots for a GGA with the PMR evolving with the organisms. This run includes the simulated annealing scheme used in the previous example and elitism of the single best fit organism. Figure 6.10 shows the evolution of the PMR for the same run. The “genetic simulated annealing” scheme is visible in the plot with the larger PMRs more efficient earlier on, and smaller PMRs dominating in the later stages. As the evolving PMR values range over nearly two orders of magnitude, it is easy to see why a single, constant choice for the PMR would be so much less efficient. Also, as one can see from the data presented, the variations in the frequency are significantly smaller than those of θ and ϕ. This idea of tailored PMRs can be extended beyond the organism, and down to the gene. Giving a separate PMR to each parameter will allow for even better adaptation. (In the natural world organisms control their mutation rates by building in DNA repair mechanisms to counteract the externally determined mutation rate set by cosmic rays and other pathogens).

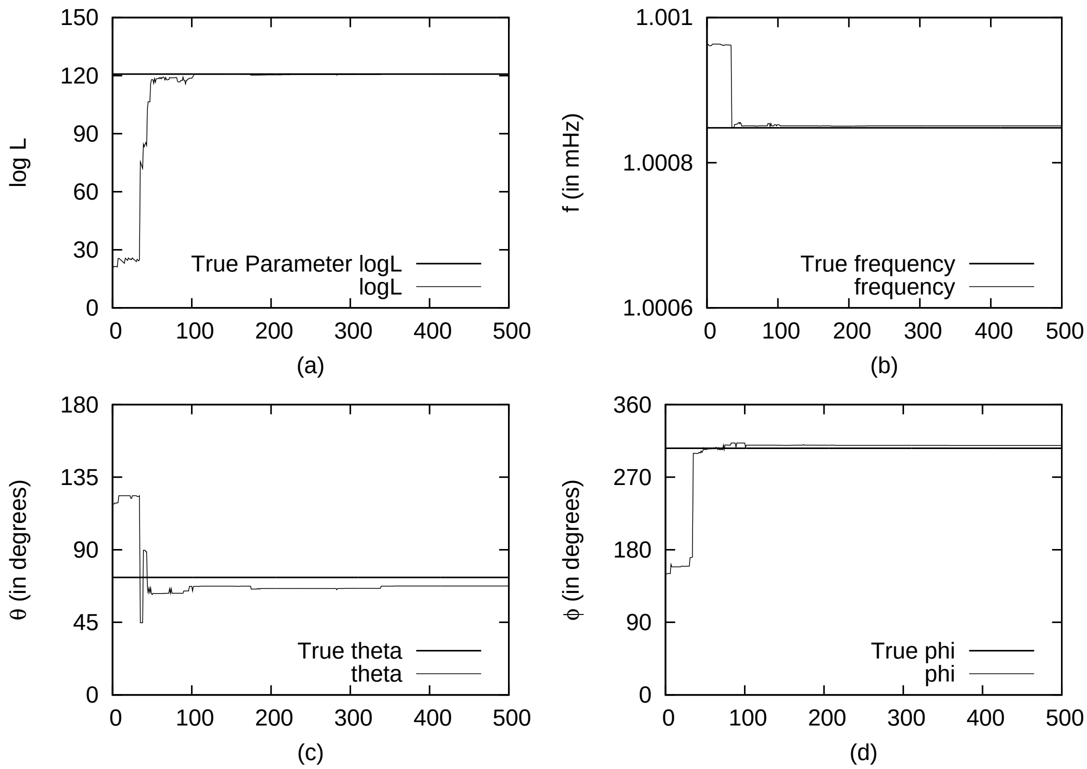

Figure 6.9: Genetic-Genetic Algorithm: Trace plots for (a) log likelihood, (b) frequency, (c), θ, and (d) ϕ for a genetic-genetic algorithm in which the PMR evolves with the organisms. The y-axes are the parameter values and log likelihoods of the best fit organism for each generation. The x-axes are generation number.

Performance in the Low SNR Regime

It is natural to question how well genetic algorithms perform in the low SNR regime. The first step is to establish a reasonable estimate for what defines the low SNR regime. This is done by applying the GGA described in the previous section to multiple realizations of a source-free data segment. A test for robustness is implemented by applying repeated GGA searches to a data stream containing a single low

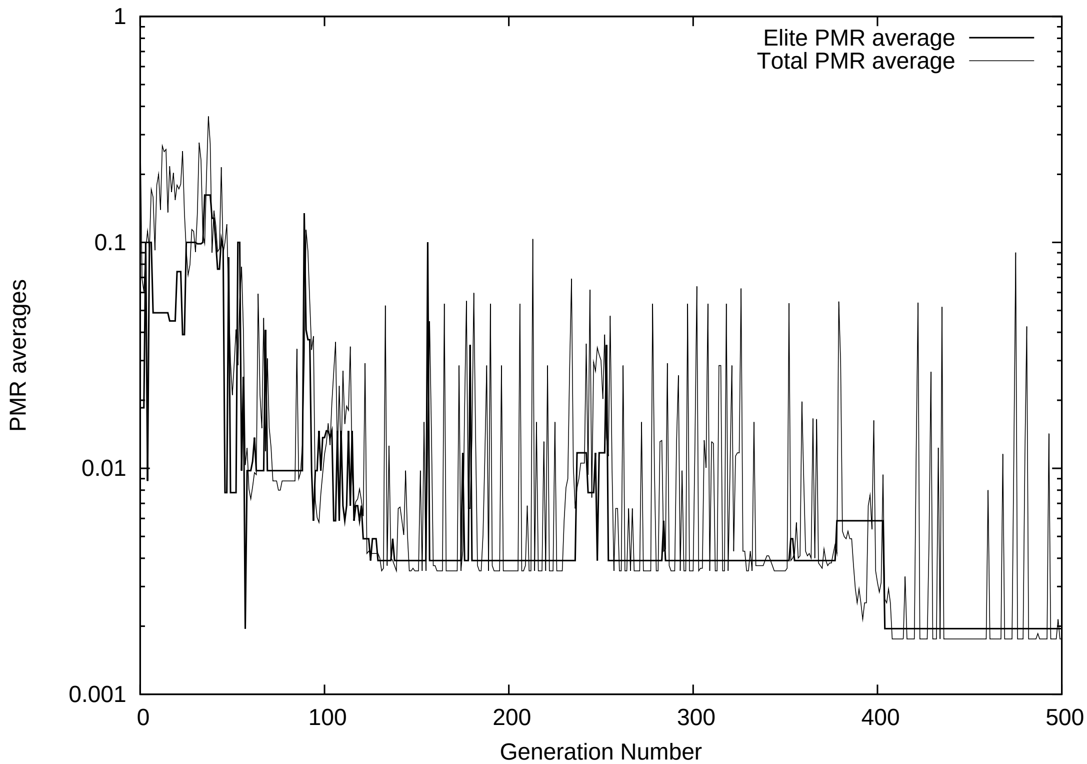

Figure 6.10: Genetic-Genetic Simulated Annealing of the PMR: Trace plots for the PMR as it evolves with the organisms. The data for this plot is from the same run that produced the data in Figure 6.9.

SNR source. Next searches of multiple realizations of low SNR, single source data streams are performed, where both the source and the noise are different in each realization. Finally, a focus is placed on a particular source that was missed by the GGA in the previous example, and the search is repeated with different noise realizations. In each case the GGA uses 10 organisms in the search.

The term "low SNR" can be defined as the SNR at which the rate of false positive and/or false negatives reach some threshold. The SNR threshold for false positives can be estimated by applying the GGA search to source-free data streams. The GGA searching for a source instead fits to the noise, and returns a false source parameter set. The SNR of the false sources returned by 1000 searches of source-free data sets were calcuated. Figure 6.11 shows the cumulative probability distribution of the recovered SNRs. A false positive rate of 2% is achieved by setting a SNR threshold of SNR > 5.9. Sources with $6 < \mathrm{SNR} < 8$ are considered to be in the low SNR regime. It should be noted that the SNR > 5.9 threshold only applies to single source searches over data segments of $100 f_m = 3.15\mu$Hz in width. Multi-source searches and searches over larger data segments will yield higher thresholds.

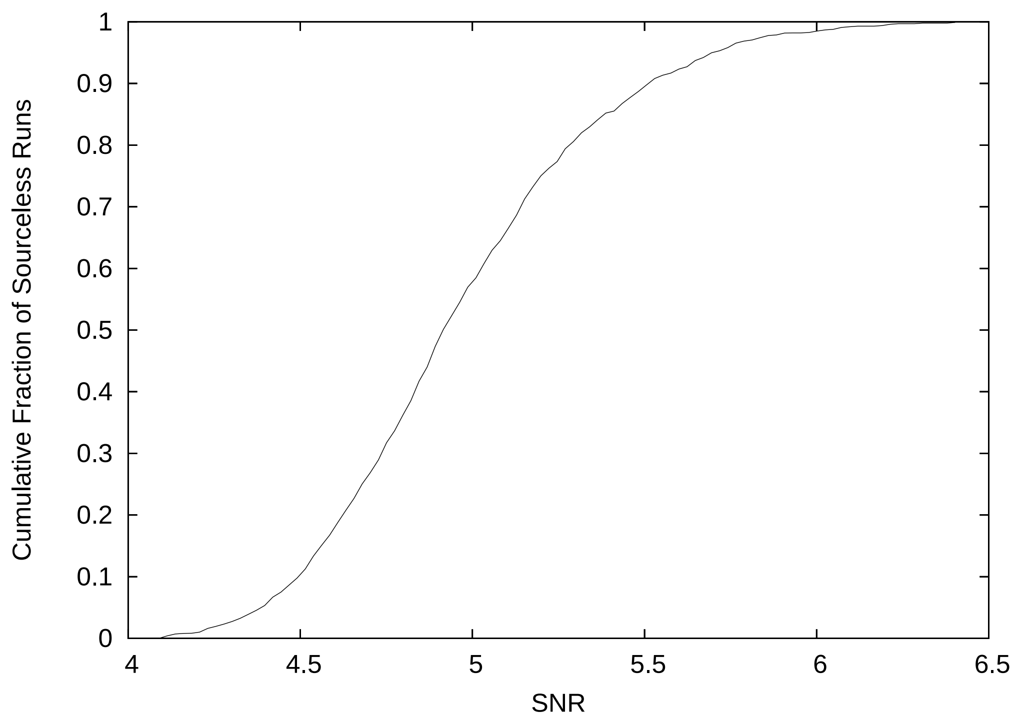

Figure 6.11: The cumulative probability distribution for the SNR of false positives found by searching source-free data.

For the repeated GGA search of a single data segment, the system used had SNR $= 7.173$ and parameters: $A = 1.19092 \times 10^{-22}$, $f = 1.003132267$ mHz, $\theta = 0.598052$, $\phi = 5.64625$, $\iota = 1.27899$, $\psi = 1.88816$, and $\gamma_o = 4.43056$. There were 100 randomly started GGA searches performed. The searches took somewhat longer than the higher SNR examples discussed earlier, but in all 100 cases the source was properly recovered.

The multiple data realization study was restricted to sources with SNRs in the range $[6.5, 7.5]$. There were 100 data realizations, and each was searched once by a GGA. Here the GGAs found the source they were seeking in 99 of the 100 cases. The failed attempt involved a source with SNR $= 6.68439$. With this SNR the expectation value of the log likelihood is 22.3, however, for the random noise realization used in the simulation the source came out to have a log likelihood of 12.57. This implies that this particular realization was very unlucky, as the noise had managed to significantly deconstructively interfere with the signal. While the GGA found the true source early in the search, it then moved off the true solution and settled on a match that was a fit to pure noise, a full $22f_m$ away from the true solution. To test this hypothesis that an unlucky noise realization had tripped up the search, another 100 searches were performed using the same source parameters, but with 100 different noise realizations. The GGA found the source in 97 of these 100 attempts. In each instance that the GGA failed to find the source it instead found a region of higher log likelihood by matching instrument noise. In that sense, the GGAs never failed in their assigned task, as any likelihood based data analysis procedure would have found the same false positives.

Multiple Sources in the Data Stream

At the low end of the LISA band there will be many thousands of sources. Thus, one expects to see multiple sources even in small segments of the data stream such as the one under consideration in this work. Simulations point to bright source densities of up to one source per two modulation frequency bins ($f_{\rm mod} = 1/\text{year}$) [48]. Thus, any search algorithm must be able to perform multiple source searches at the low end of the LISA band.

Figure 6.12 shows an implementation of the GGA with standard simulated annealing to a LISA data stream snippet of width $100f_{\rm mod}$, containing five monochromatic binary systems. The standard simulated annealing was completed in the first

$g_{\text{cool}} = 4000$ generations, by which time the GGA had separated out the values for the source frequencies and co-latitudes. The grouping of azimuthal angles was separated soon thereafter, with minor modifications of the parameters occurring over the next 5000 generations. Search results are summarized in Table 6.2. The GGA accurately recovered the source parameters in this and similar multiple $(3-5)$ source data sets, converging to a best fit solution in less than 5000 generations per source with 10 organisms per generation, so long as the source correlation coefficients were below ~ 0.25. The intrinsic parameters for the sources were recovered to within 2σ of the true parameters (based on a Fisher Information Matrix estimate of the uncertainties of the recovered parameters). When highly correlated sources are used, the GGA spends a correspondingly longer time to pick out the source parameters. Investigations in this area were limited. A full study of the affect of source correlation on computational cost is to be carried out in the future.

Table 6.2: GGA search for 5 galactic binaries. The frequencies are quoted relative to 1 mHz as $f = 1\,\text{mHz} + \delta f$ with δf in μHz. All angles are quoted in radians.

	SNR	A (10^{-22})	δf	θ	ϕ	ψ	ι	φ_0
True	12.7	1.02	1.638	2.77	1.48	2.28	0.886	0.273
GA ML	11.6	1.08	1.635	2.86	1.40	2.63	1.02	5.94
True	19.3	2.23	0.7000	2.41	5.87	0.435	1.88	4.29
GA ML	17.7	2.11	0.7008	2.43	5.90	0.460	1.86	4.20
True	17.8	1.74	0.3937	0.756	1.85	1.41	2.02	3.09
GA ML	17.0	1.80	0.3942	0.777	1.84	1.27	1.95	2.57
True	15.8	2.16	1.002	1.53	1.30	1.35	1.70	4.63
GA ML	14.8	2.17	1.002	1.59	1.28	1.37	1.68	4.68
True	12.1	0.836	1.944	0.872	0.802	1.56	0.805	3.87
GA ML	11.8	1.09	1.950	0.876	0.803	2.87	1.09	3.48

Using Active Organisms

So far all of the organisms that have been discussed are passive organisms. They are passive in the sense that once they are bred, the organisms themselves remain

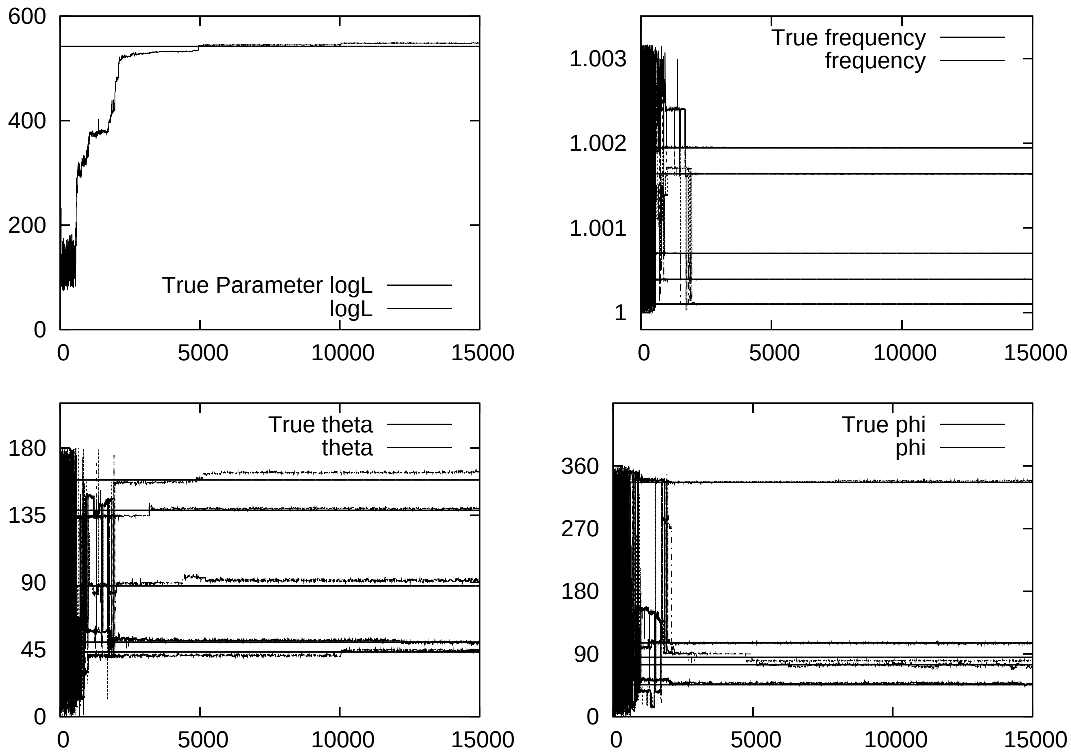

Figure 6.12: Genetic algorithm search for 5 sources: Trace plots for (a) log likelihood, (b) frequency, (c), θ, and (d) ϕ for a genetic algorithm searching for the presence of five gravitational wave sources in the data stream. The y-axes are the parameter values and log likelihoods of the best fit organism for each generation. The x-axes are generation number.

unchanged, and are simply used to breed the next generation. One can imagine organisms that 'learn' during their lifetime, advancing toward a better solution. Directed search methods such as an uphill simplex, i.e. an amoeba, provide a means for organisms to advance within a generation. As the likelihood surface is not entirely smooth, the simplex may get stuck in a local maximum that is removed from the global maximum. So the generational process is still necessary to ensure full exploration of the surface. One approach is to use the the parameters bred from one generation as the centroid of the simplex (amoeba), which will then proceed to move uphill across the likelihood surface. Another approach, that will be described in a future publication, is to use 'Genetic Amoeba', where genes code for each vertex of the simplex. The amoeba are allowed to breed after they have found enough food (i.e. increased their likelihood by a specified amount). Amoeba that eat well get to breed the most often and have the most offspring.

Figure 6.13 shows trace plots for an implementation of a GGA with a single directed organism per generation. The other 9 organisms were the standard passive organisms. There was elitism with a single organism being cloned into the succeeding generation, and there was no standard simulated annealing. What is missing from the plot is the computational cost. While computational cost can easily be derived from the plots with passive organisms, active organisms, such as an uphill simplex involve multiple calls to the F-statistic function within a single generation. At the 8^{th} generation, where the search surpasses the true likelihood value, the computational cost is $\$ = 876$. This cost is slightly lower than the cost of a GGA with only passive organisms at the point where its search surpasses the likelihood value for the true parameters. However, for true LISA data, we will not know the true parameters, and thus will have to allow the algorithms to undergo extended runs to ensure they have fully explored the space and found the global maximum. The higher computational cost per generation of the simplex method (which averages $\sim$ 100 calls to find a local maximum) will quickly lead to a higher total cost of the search. Other directed methods that are more efficient than an uphill simplex may provide an alternative that will provide an overall improvement in efficiency. Future work will include an examination of other possibly more efficient directed methods, and a detailed study of the Genetic Amoeba algorithm.

Future Plans

The next step is to investigate the limits of the algorithm both in terms of source

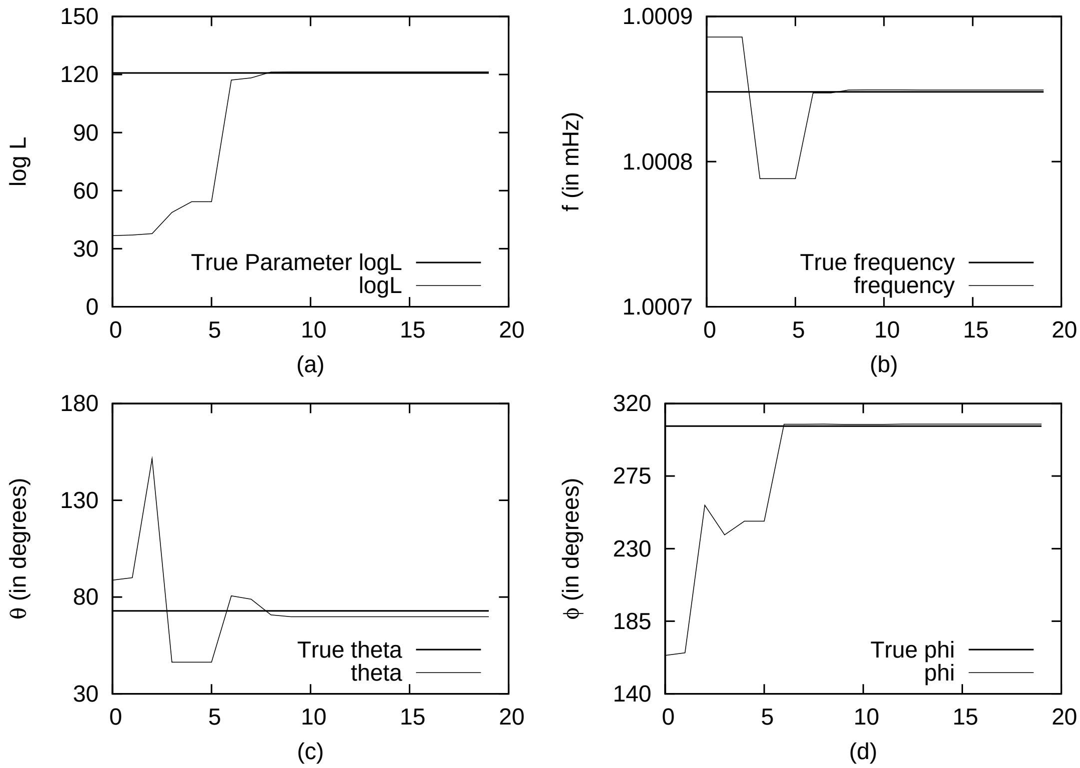

Figure 6.13: GGA with a directed organism: Trace plots for (a) log likelihood, (b) frequency, (c), θ, and (d) ϕ for a GGA with a single directed organism. The y-axes are the parameter values and log likelihoods of the best fit organism for each generation. The x-axes are generation number.

number and source density across the low frequency regime of the LISA band. While an optimal solution would employ a matched filter that includes every resolvable source in the LISA band [54], it is unlikely that a direct search for this "super template" is the best way to proceed. A better approach may be to start with a collection of "single cell" organism that each code for a single source (or possibly small collections of highly correlated sources), then combine these cells into a multi-cellular organism that searches for the super template. This approach is motivated by the *cellular slime molds* Dictyostelida and Acrasida, which spend most of their lives as separate single-celled amoeboid protists, but upon the release of a chemical signal, the individual cells aggregate into a great swarm that acts as a single multicelluar organism, capable of movement and the formation of large fruiting bodies. Future work will also include investigations into algorithm optimization and adaptation of the algorithm to other source types (e.g. coalescing binaries). Furthermore, a thorough study comparing the computational cost and resolution capabilities of an optimized genetic algorithm to other (optimized) search methods like Markov Chain Monte Carlo searches, gClean, Slice & Dice, and Maximum Entropy methods would provide guidance on how to proceed in solving the LISA Data Analysis Challenge.

CHAPTER 7

MARKOV CHAIN MONTE CARLO: SEARCHING AND STATISTICS

Introduction

This chapter focuses on the use of MCMC algorithms as methods for searching the LISA data stream to determine the source parameters of the monochromatic binary systems used to create the simulated LISA signal, and for providing an estimate of the posterior distributions of those recovered parameters. The bulk of this work was first published in Ref. [54].

MCMC methods have been used to study the extraction of coalescing binary [87] and spinning NS [88] signals from terrestrial interferometers. Recently, MCMC methods have been applied to a simplified toy problem [89] that shares some of the features of the LISA Cocktail Party Problem described in Chapter 3. These studies have shown that MCMC methods hold considerable promise for gravitational wave data analysis, and offer many advantages over the standard template grid searches. For example, the EMRI data analysis problem [73, 78] is often cited as the greatest challenge facing LISA science. Neglecting the spin of the smaller body yields a 14 dimensional parameter space, which would require $\sim 10^{40}$ templates to explore in a grid based search [78]. This huge computational cost arises because grid based searches scale geometrically with the parameter space dimension D. In contrast, the computational cost of MCMC based searches scale linearly with D. In fields such as finance, MCMC methods are routinely applied to problems with $D > 1000$, making the LISA EMRI problem seem trivial in comparison.

In the case of compact stellar binaries [80, 81, 82, 83, 84] and EMRIs [73, 78], the number of sources is likely to be so large that it will be impossible to resolve all the sources individually, so that there will be a residual signal that is variously referred to as a confusion limited background or confusion noise. It is important that this confusion noise be made as small as possible so as not to hinder the detection of other high value targets. Several estimates of the confusion noise level have been made [48, 60, 80, 81, 84, 85], and they all suggest that unresolved signals will be the dominant source of low frequency noise for LISA. However, these estimates are based on assumptions about the efficacy of the data analysis algorithms that will be used to identify and regress sources from the LISA data stream, and it is unclear at present how reasonable these assumptions might be. The upshot is that the LISA data stream will contain the signals from tens of thousands of individual sources, and ways must be found to isolate individual voices from the crowd. This Cocktail Party Problem is the central issue in LISA data analysis.

As described in Chapter 3, optimal filtering of the LISA data would require the construction of a filter bank that described the signals from every source that contributes to the data stream. In principle one could construct a vast template bank describing all possible sources and look for the best match with the data. In practice the enormous size of the search space and the presence of unmodeled sources renders this direct approach impractical. Possible alternatives to a full template based search include iterative refinement of a source-by-source search, ergodic exploration of the parameter space using MCMC algorithms , Darwinian optimization by genetic algorithms similar to those discussed in Chapter 6, and global iterative refinement using the Maximum Entropy Method (MEM). Each approach has its strengths and weakness, and at this stage it is not obvious which approach will prove superior. In this work the popular MCMC [90, 91] method is applied to simulated LISA data. The simulated data streams contain the signals from multiple galactic binaries.

After the executive summary of results, the structure of this chapter follows the development sequence taken to arrive at a fast and robust MCMC algorithm. A basic MCMC algorithm is introduced and then applied to a full 7 parameter search for a single galactic binary. The performance of a basic MCMC algorithm that uses the F-statistic is then studied and a number of problems with this simple approach are identified. A more advanced mixed MCMC algorithm that incorporates simulated annealing is then introduced and is successfully applied to multi-source searches. The issue of model selection is then addressed and approximate Bayes factor are calculated by super-cooling the Markov Chains to extract maximum likelihood estimates. At the end of the chapter is a discussion of future refinements and extensions of this approach.

Chapter Summary

This application of the MCMC method to LISA data analysis will demonstrate the method to have considerable promise. This work confirms that an F-statistic based MCMC algorithm can handle a source density of one source per ten frequency bins across a one hundred bin snippet. Some evidence will be seen that high local source densities pose a challenge to the latest algorithm. The lesson, as will be shown, is that adding new, specially tailored proposal distributions to the mix helps to keep the chain from sticking at secondary modes of the posterior (it takes a cocktail to solve the Cocktail Party Problem). On the other hand, evidence will indicate the presence of strong multi-modalities whereby the secondary modes have likelihoods within a few percent of the global maximum. In those cases the chain will tend to jump back and forth between modes before being forced into a decision by the super-cooling process that follows the main MCMC run. Indeed, these algorithms may already be pushing the limits of what is possible using any data analysis method. For example, a 10 source search will use a model with 70 parameters to fit 400 pieces of data (2 channels $\times$ 2 Fourier components $\times$ 100 bins). One of the goals is to better understand the theoretical limits of what can be achieved so that it is known when to stop trying to improve the algorithm!

Markov Chain Monte Carlo

These investigations began by implementing a basic MCMC search for galactic binaries that spans the full $D = 7N$ dimensional parameter space using the Metropolis-Hastings [91] algorithm. The idea is to generate a set of samples, $\{\vec{x}\}$, that correspond to draws from the posterior distribution, $p(\vec{\lambda}|s)$. To do this one starts at a randomly chosen point $\vec{x}$ and generates a Markov chain according to the following algorithm: Using a proposal distribution $q(\cdot|\vec{x})$, draw a new point $\vec{y}$. Evaluate the Hastings ratio

$$H = \frac{p(\vec{y})p(s|\vec{y})q(\vec{x}|\vec{y})}{p(\vec{x})p(s|\vec{x})q(\vec{y}|\vec{x})} \,. \tag{7.1}$$

Accept the candidate point $\vec{y}$ with probability $\alpha = \min(1, H)$, otherwise remain at the current state $\vec{x}$ (Metropolis rejection [90]). Remarkably, this sampling scheme produces a Markov chain with a stationary distribution equal to the posterior distribution of interest, $p(\vec{\lambda}|s)$, regardless of the choice of proposal distribution [52]. A concise introduction to MCMC methods can be found in the review paper by Andrieu *et al* [92].

On the other hand, a poor choice of the proposal distribution will result in the algorithm taking a very long time to converge to the stationary distribution (known as the burn-in time). Elements of the Markov chain produced during the burn-in phase have to be discarded as they do not represent the stationary distribution. When dealing with large parameter spaces the burn-in time can be very long if poor techniques are used. For example, the Metropolis sampler, which uses symmetric proposal distributions, explores the parameter space with an efficiency of at most $\sim 0.3/D$, making it a poor choice for high dimension searches. Regardless of the sampling scheme, the mixing of the Markov chain can be inhibited by the presence of strongly correlated parameters. Correlated parameters can be dealt with by making a local coordinate transformation at $\vec{x}$ to a new set of coordinates that diagonalises the FIM, $\Gamma_{ij}(\vec{x})$.

A number of proposal distributions and update schemes to search for a single galactic binary were tried. The results were very disappointing. Bold proposals that attempted large jumps had a very poor acceptance rate, while timid proposals that attempted small jumps had a good acceptance rate, but they explored the parameter space very slowly, and got stuck at local modes of the posterior. Lorentzian proposal distributions fared the best as their heavy tails and concentrated peaks lead to a mixture of bold and timid jumps, but the burn in times were still very long and the subsequent mixing of the chain was torpid. The MCMC literature is full of similar examples of slow exploration of large parameter spaces, and a host of schemes have been suggested to speed up the burn-in. Many of the accelerated algorithms use adaptation to tune the proposal distribution. This violates the Markov nature of the chain as the updates depend on the history of the chain. More complicated adaptive algorithms have been invented that restore the Markov property by using additional Metropolis rejection steps. The popular Delayed Rejection Method [93] and Reversible Jump Method [94] are examples of adaptive MCMC algorithms. A simpler approach is to use a non-Markov scheme during burn-in, such as adaptation or simulated annealing, then transition to a Markov scheme after burn-in. Since the burn-in portion of the chain is discarded, it does not matter if the MCMC rules are broken (the burn-in phase is more like Las Vegas than Monte Carlo).

Before resorting to complex acceleration schemes a much simpler approach was applied that proved to be very successful. When using the Metropolis-Hastings algorithm there is no reason to restrict the updates to a single proposal distribution. For example, every update could use a different proposal distribution so long as the choice of distribution is not based on the history of the chain. The proposal distributions to be used at each update can be chosen at random, or they can be applied in a fixed sequence. The experience gained with single proposal distributions suggested that a scheme that combined a very bold proposal with a very timid proposal would lead to fast burn-in and efficient mixing. For the bold proposal a uniform distribution was chosen for each of the source parameters $\vec{\lambda} \rightarrow (A, f, \theta, \phi, \psi, \iota, \varphi_0)$. Here A is the amplitude, f is the gravitational wave frequency, θ and ϕ are the ecliptic co-latitude and longitude, ψ is the polarization angle, ι is the inclination of the orbital plane, and φ_0 is the orbital phase at some fiducial time. The amplitudes were restricted to the range $A \in [10^{-23}, 10^{-21}]$ and the frequencies were restricted to lie within the range of the data snippet $f \in [0.999995, 1.003164]$ mHz (the data snippet contained 100 frequency bins of width $\Delta f = 1/\text{year}$). A better choice would have been to use a cosine distribution for the co-latitude θ and inclination ι, but the choice is not particularly important. When multiple sources were present each source was updated separately during the bold proposal stage. For the timid proposal a normal distribution was used for each eigendirection of the FIM, $\Gamma_{ij}(\vec{x})$. The standard deviation $\sigma_{\hat{k}}$ for each eigendirection k was set equal to $\sigma_{\hat{k}} = 1/\sqrt{\alpha_{\hat{k}} D}$, where $\alpha_{\hat{k}}$ is the corresponding eigenvalue of $\Gamma_{ij}(\vec{x})$, and $D = 7N$ is the search dimension. The factor of $1/\sqrt{D}$ ensures a healthy acceptance rate as the typical total jump is then $\sim 1\sigma$. All N sources were updated simultaneously during the timid proposal stage. Note that the timid proposal distributions are not symmetric since $\Gamma_{ij}(\vec{x}) \neq \Gamma_{ij}(\vec{y})$. One set of bold proposals (one for each source) was followed by ten timid proposals in a repeating cycle. The ratio

Table 7.1: 7 parameter MCMC search for a single galactic binary

	A (10^{-22})	f (mHz)	θ	ϕ	ψ	ι	φ_0
$\vec{\lambda}_{\rm True}$	1.73	1.0005853	0.98	4.46	2.55	1.47	0.12
$\vec{\lambda}_{\rm MCMC}$	1.44	1.0005837	1.07	4.42	2.56	1.52	0.15
$\sigma_{\rm FIM}$	0.14	2.2e-06	0.085	0.051	0.054	0.050	0.22
$\sigma_{\rm MCMC}$	0.14	2.4e-06	0.089	0.055	0.058	0.052	0.23

of the number of bold to timid proposals impacted the burn-in times and the final mixing rate, but ratios anywhere from 1:1 to 1:100 worked well. Uniform priors were used, $p(\vec{x}) = $ const., for all the parameters, though once again a cosine distribution would have been better for θ and ι. Two independent LISA data channels were simulated directly in the frequency domain using the method described in Ref. [48], with the sources chosen at random using the same uniform distributions employed by the bold proposal. The data covers 1 year of observations, and the data snippet contains 100 frequency bins (of width 1/year). The instrument noise was assumed to be stationary and Gaussian, with position noise spectral density $S_n^{\rm pos} = 4 \times 10^{-22}$ m^2 Hz^{-1} and acceleration noise spectral density $S_n^{\rm accel} = 9 \times 10^{-30}$ m^2 s^{-4} Hz^{-1}.

Table 7.1 summarizes the results of one MCMC run using a model with one source to search for a single source in the data snippet. Burn-in lasted ~ 2000 iterations, and post burn-in the chain was run for 10^6 iterations with a proposal acceptance rate of 77% (the full run took 20 minutes on a Mac G5 2 GHz processor). The chain was used to calculate means and variances for all the parameters. The parameter uncertainty estimates extracted from the MCMC output are compared to the FIM estimates evaluated at the mean values of the parameters. The source had true SNR $= 12.9$, and MCMC recovered SNR $= 10.7$. Histograms of the posterior parameter distributions are shown in Figure 7.1, where they are compared to the Gaussian approximation to the posterior given by the FIM. The agreement is impressive, especially considering that the bandwidth of the source is roughly 10 frequency bins, so there are very few noise samples to work with. Similar results were found for other MCMC runs on the same source, and for MCMC runs with other sources. Typical burn-in times were of order 3000 iterations, and the proposal acceptance rate was around 75%.

The algorithm was run successfully on two and three source searches (the model dimension was chosen to match the number of sources in each instance), but on occasions the chain would get stuck at a local mode of the posterior for a large number of iterations. Before attempting to cure this problem with a more refined MCMC algorithm, it was decided to eliminate the extrinsic parameters $A, \iota, \psi, \varphi_0$ from the search by using a multi-filter generalized F-statistic. This reduces the search dimension to $D = 3N$, with the added benefit that the projection onto the (f, θ, ϕ) sub-space yields a softer target for the MCMC search.

F-Statistic MCMC

An F-statistic based MCMC algorithm was implemented using the F-statistic approach described in Chapter 3, but with the full likelihood replaced by the F-statistic and the full FIM replaced by the projected FIM. Applying the F-MCMC search to the same data set as before yields the results summarized in Figure 7.2 and Table 7.2. The recovered source parameters and signal-to-noise ratio (SNR = 10.4) are very similar to those found using the full 7-parameter search, but the F-MCMC estimates for the errors in the extrinsic parameters are very different. This is because the chain does not explore extrinsic parameters, but rather relies upon the F-statistic to find the extrinsic parameters that give the largest log likelihood based on the current values for the intrinsic parameters. The effect is very pronounced in the histograms shown in Figure 7.2. Similar results were found for other F-MCMC runs on the same source, and for F-MCMC runs with other sources. Typical burn-in times

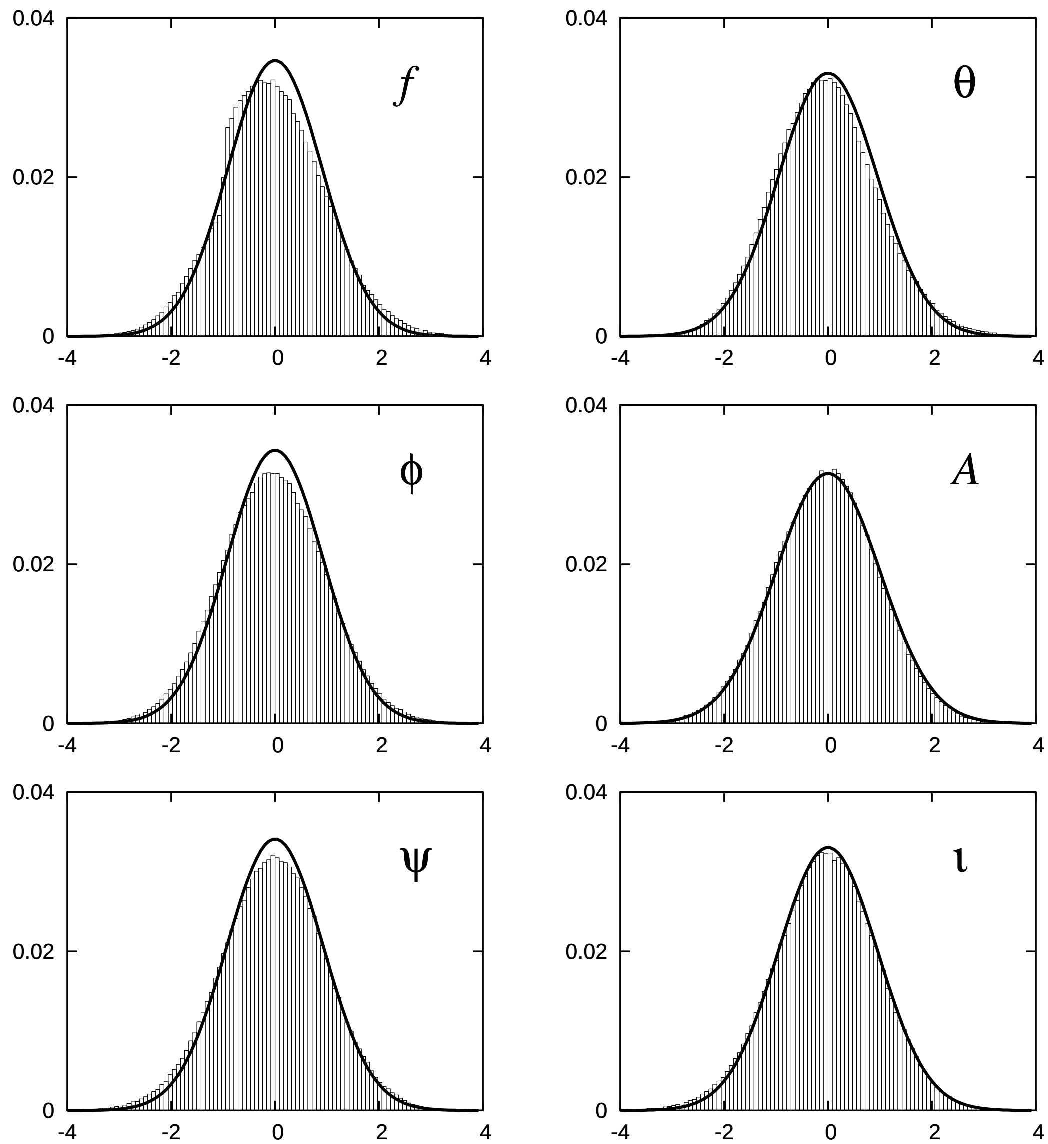

Figure 7.1: Histograms showing the posterior distribution (grey) of the parameters. Also shown (black line) is the Gaussian approximation to the posterior distribution based on the FIM. The mean values have been subtracted, and the parameters have been scaled by the square root of the variances calculated from the MCMC chains.

Table 7.2: F-MCMC search for a single galactic binary

	A (10^{-22})	f (mHz)	θ	ϕ	ψ	ι	φ_0
$\vec{\lambda}_{\rm True}$	1.73	1.0005853	0.98	4.46	2.55	1.47	0.12
$\vec{\lambda}_{\rm F-MCMC}$	1.38	1.0005835	1.09	4.42	2.56	1.51	0.17
$\sigma_{\rm FIM}$	0.14	2.2e-06	0.089	0.052	0.055	0.051	0.22
$\sigma_{\rm MCMC}$	0.02	2.5e-06	0.093	0.056	0.027	0.016	0.21

were of order 1000 iterations, and the proposal acceptance rate was around 60%. As expected, the F-MCMC algorithm gave shorter burn-in times than the full parameter MCMC, and a comparable mixing rate.

Of note is that since the search is now over the projected sub-space $\{f_J, \theta_J, \phi_J\}$ of the full parameter space, and the full FIM, $\Gamma_{ij}(\vec{x})$, was replaced by the projected FIM, $\gamma_{ij}(\vec{x})$, the projection of the $k^{\rm th}$ parameter is given by

$$\Gamma^{n-1}_{ij} = \Gamma^{n}_{ij} - \frac{\Gamma^{n}_{ik}\Gamma^{n}_{jk}}{\Gamma^{n}_{kk}}\,, \tag{7.2}$$

where n denotes the dimension of the projected matrix. Repeated application of the above projection yields $\gamma_{ij} = \Gamma^{3N}_{ij}$. Inverting γ_{ij} yields the same uncertainty estimates for the intrinsic parameters as one gets from the full FIM, but the covariances are much larger. The large covariances make it imperative that the proposal distributions use the eigenvalues and eigenvectors of γ_{ij}, as using the parameter directions themselves would lead to a slowly mixing chain.

It is interesting to compare the computational cost of the F-MCMC search to a traditional F-Statistic based search on a uniformly spaced template grid. To cover the parameter space of one source (which for the current example extends over the full sky and 100 frequency bins) with a minimal match [37] of MM $= 0.9$ requires 39,000 templates [95]. A typical F-MCMC run uses less than 1000 templates to cover the same search space. The comparison becomes even more lopsided if one considers

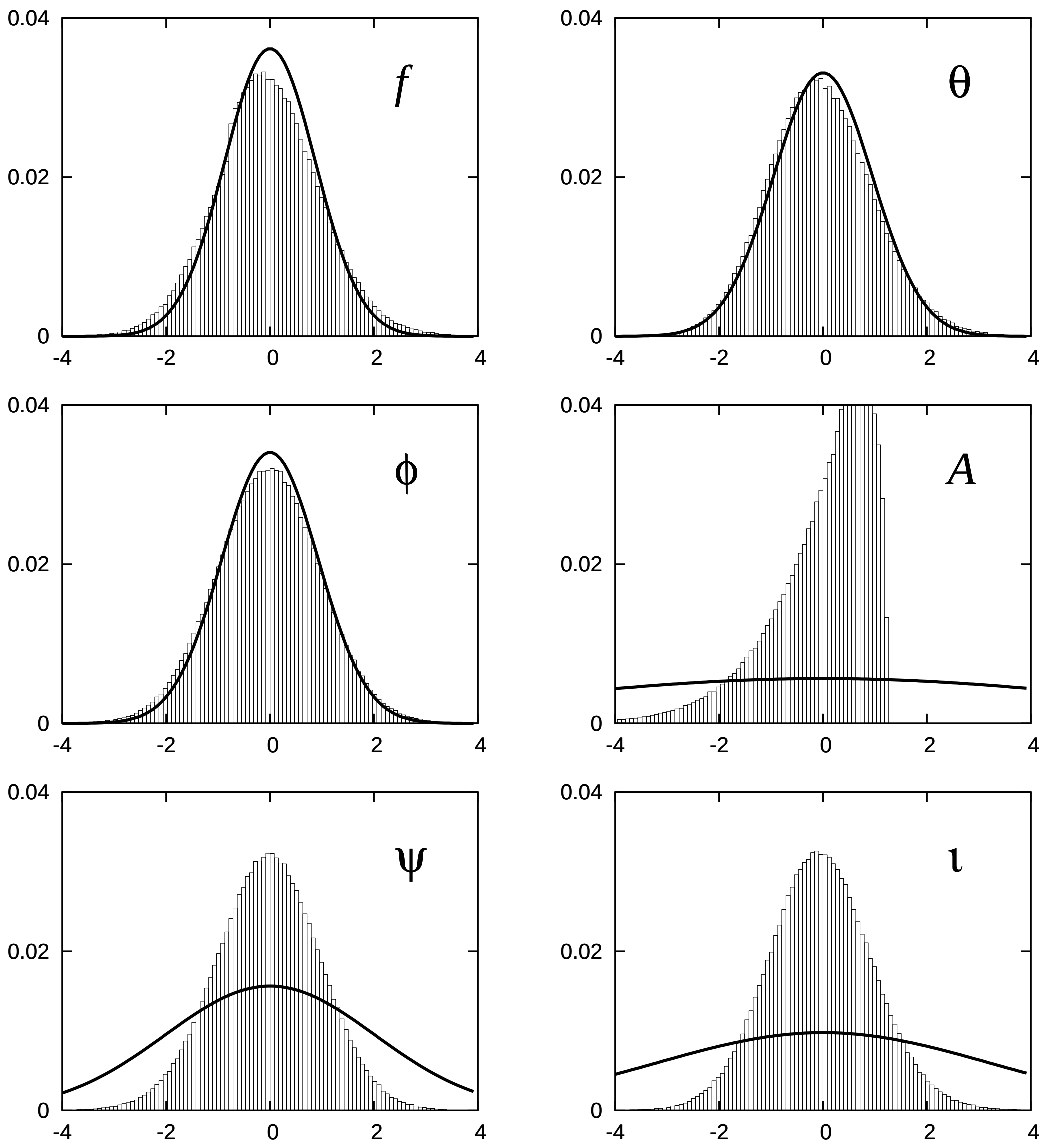

Figure 7.2: Histograms showing the posterior distribution (grey) of the parameters. Also shown (black line) is the Gaussian approximation to the posterior distribution based on the FIM. The mean values have been subtracted, and the parameters have been scaled by the square root of the variances calculated from the F-MCMC chains.

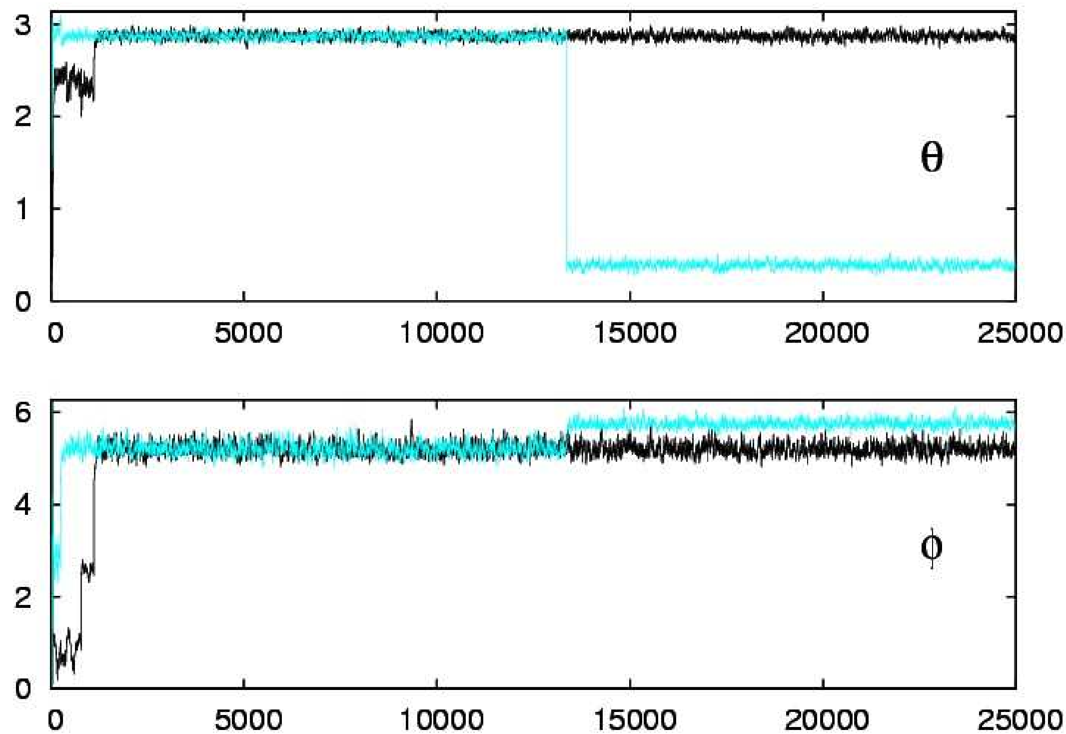

Figure 7.3: Trace plots of the sky location parameters for two F-MCMC runs on the same data set. Both chains initially locked onto a secondary mode of the posterior, but one of the chains (light colored line) transitioned to the correct mode after 13,000 iterations.

simultaneous searches for multiple sources. A grid based simultaneous search for two sources using the F-statistic would take $(39,000)^2 \simeq 1.5 \times 10^9$ templates, while the basic F-MCMC algorithm typically converges on the two sources in just 2000 steps. As the number of sources in the model increases the computation cost of the grid based search grows geometrically while the cost of the F-MCMC search grows linearly. It is hard to imagine a scenario (other than quantum computers) where non-iterative grid based searches could play a role in LISA data analysis.

While testing the F-MCMC algorithm on different sources instances were discovered where the chain became stuck at secondary modes of the posterior. A good example occurred for a source with parameters $(A, f, \theta, \phi, \psi, \iota, \varphi_0)$=(1.4e-22, 1.0020802 mHz,

$0.399, 5.71, 1.3, 0.96, 1.0)$ and SNR $= 16.09$. Most MCMC runs returned good fits to the source parameters, with an average log likelihood of $\ln \mathcal{L} = 132$, mean intrinsic parameter values $(f, \theta, \phi) = (1.0020809\,\text{mHz}, 0.391, 5.75)$ and SNR $= 16.26$. However, some runs locked into a secondary mode with average log likelihood $\ln \mathcal{L} = 100$, mean intrinsic parameter values $(f, \theta, \phi) = (1.0020858\,\text{mHz}, 2.876, 5.20)$ and SNR $= 14.15$. It could sometimes take hundreds of thousands of iterations for the chain to discover the dominant mode. Figure 7.4 shows plots of the (inverted) likelihood $\mathcal{L}$ and the log likelihood $\ln \mathcal{L}$ as a function of sky location for fixed $f = 1.0020802\,\text{mHz}$. The log likelihood plot reveals the problematic secondary mode near the south pole, while the likelihood plot shows just how small a target the dominant mode presents to the F-MCMC search. Similar problems with secondary modes were encountered in the $f-\phi$ plane, where the chain would get stuck a full bin away from the correct frequency. These problems with the basic F-MCMC algorithm motivated the embellishments described in the following section.

Multiple Proposals and Heating

The LISA data analysis problem belongs to a particularly challenging class of MCMC problems known as "mixture models." As the name suggests, a mixture model contains a number of components, some or all of which may be of the same type. In this present study all the components are slowly evolving, circular binaries, and each component is described by the same set of seven parameters. There is nothing to stop two components in the search model from latching on to the same source, nor is there anything to stop one component in the search model from latching on to a blend of two overlapping sources. In the former instance the likelihood is little improved by using two components to model one source, so over time one of the components will tend to wander off in search of another source. In the latter instance

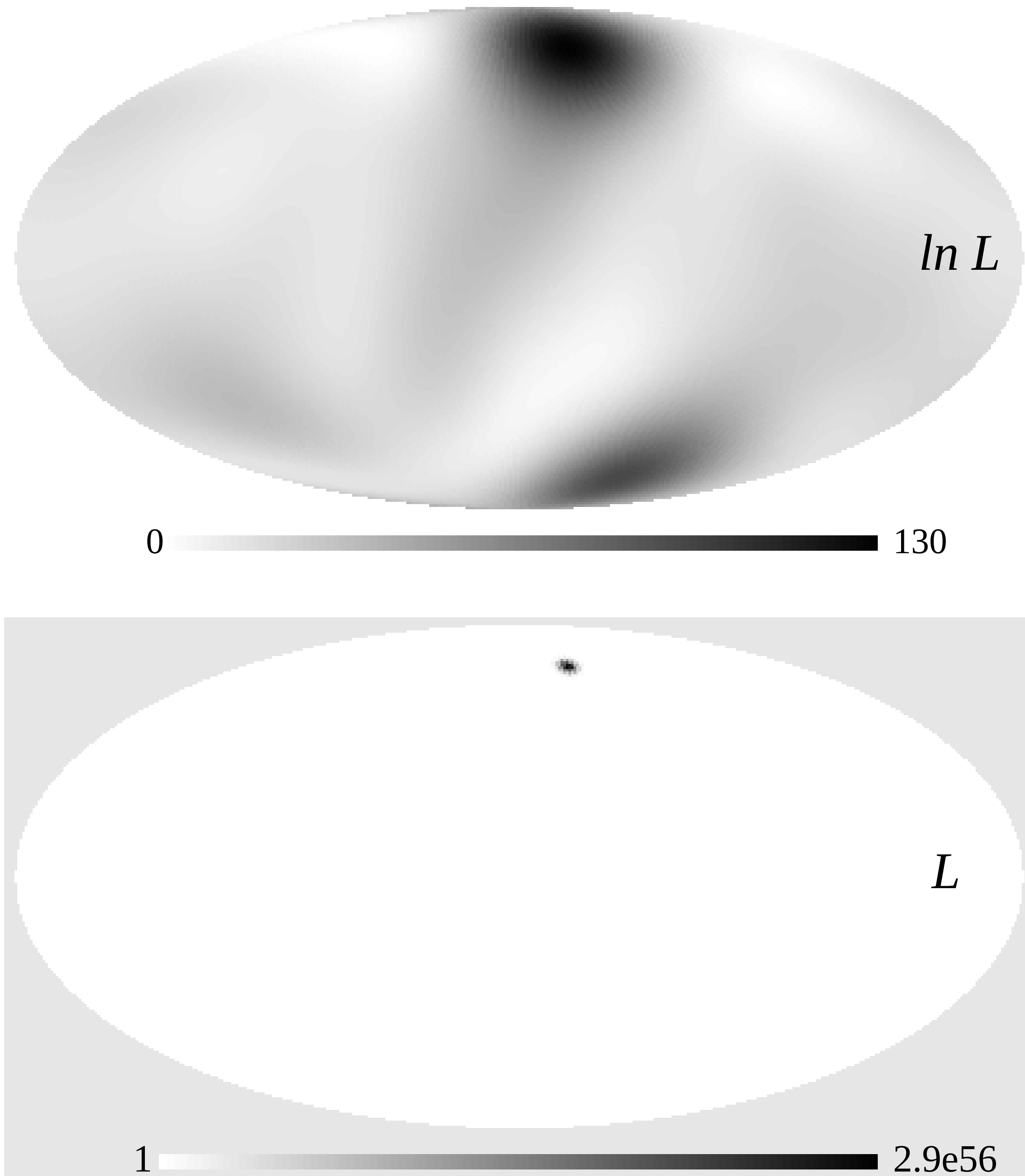

Figure 7.4: Inverted likelihood and log likelihood as a function of sky location at fixed frequency.

it may prove impossible for any data analysis method to de-blend the sources (the marginal likelihood for the single component fit to the blended sources may exceed the marginal likelihood of the "correct" solution).

The difficulties encountered with the single source searches getting stuck at secondary modes of the posterior are exacerbated in the multi-source case. Source overlaps can create additional secondary modes that are not present in the non-overlapping case. Two techniques were empolyed to speed burn-in and to reduce the chance of the chain getting stuck at a secondary mode: simulated annealing and multiple proposal distributions. Simulated annealing works by softening the likelihood function, making it easier for the chain to move between modes. The likelihood (3.4) can be thought of as a partition function $Z = C\exp(-\beta E)$ with the "energy" of the system given by $E = (s - h|s - h)$ and the "inverse temperature" equal to $\beta = 1/2$. The goal is to find the template h that minimizes the energy of the system. Heating up the system by setting $\beta < 1/2$ allows the Markov Chain to rapidly explore the likelihood surface. A standard power law cooling schedule was used,

$$\beta = \begin{cases} \beta_0 \left(\frac{1}{2\beta_0}\right)^{t/T_c} & 0 < t < T_c \\ \frac{1}{2} & t \geq T_c \end{cases} \tag{7.3}$$

where t is the number of steps in the chain, T_c is the cooling time and β_0 is the initial inverse temperature. It took some trial and error to find good values of T_c and β_0. If some of the sources have very high SNR it is a good idea to start at a high temperature $\beta_0 \sim 1/50$, but in most cases it was found $\beta_0 = 1/10$ to be sufficient. The optimal choice for the cooling time depends on the number of sources and the initial temperature. It was necessary to increase T_c roughly linearly with the the number of sources and the initial temperature. Setting $T_c = 10^5$ for a model with $N = 10$ sources and an initial temperature of $\beta_0 = 1/10$ gave fairly reliable results, but it is always a good idea to allow longer cooling times if the computational resources are available. The portion of the chain generated during the annealing phase has to be discarded as the cooling introduces an arrow of time which necessarily violates the reversibility requirement of a Markov Chain.

After cooling to $\beta = 1/2$ the chain can explore the likelihood surface for the purpose of extracting parameter estimates and error estimates. Finally, one can extract maximum likelihood estimates by "super cooling" the chain to some very low temperature ($\beta \sim 10^4$ was used here).

The second ingredient in the advanced F-MCMC algorithm is a large variety of proposal distributions. The following types of proposal distribution were used: Uniform$(\cdot, \vec{x}, i)$ - a uniform draw on all the parameters that describe source i, using the full parameter ranges, with all other sources held fixed; Normal$(\cdot, \vec{x})$ - a multivariate normal distribution with variance-covariance matrix given by $3N \times \gamma(\vec{x})$; Sky$(\cdot, \vec{x}, i)$ - a uniform draw on the sky location for source i; σ-Uniform$(\cdot, \vec{x}, i)$ - a uniform draw on all the parameters that describe source i, using a parameter range given by some multiple of the standard deviations given by $\gamma(\vec{x})$. The Uniform$(\cdot, \vec{x}, i)$ and Normal$(\cdot, \vec{x})$ proposal distributions are the same as those used in the basic F-MCMC algorithm. The Sky$(\cdot, \vec{x}, i)$ proposal proved to be very useful at getting the chain away from secondary modes like the one seen in Figure 7.4, while the σ-Uniform$(\cdot, \vec{x}, i)$ proposal helped to move the chain from secondary modes in the $f - \phi$ or $f - \theta$ planes. During the initial annealing phase the various proposal distributions were used in a cycle with one set of the bold distributions (Uniform, Sky and $\sigma-$Uniform) for every 10 draws from the timid multivariate normal distribution. During the main MCMC run at $\beta = 1/2$ the ratio of timid to bold proposals was increased by a factor of 10, and in the final super-cooling phase only the timid multivariate normal distribution was used.

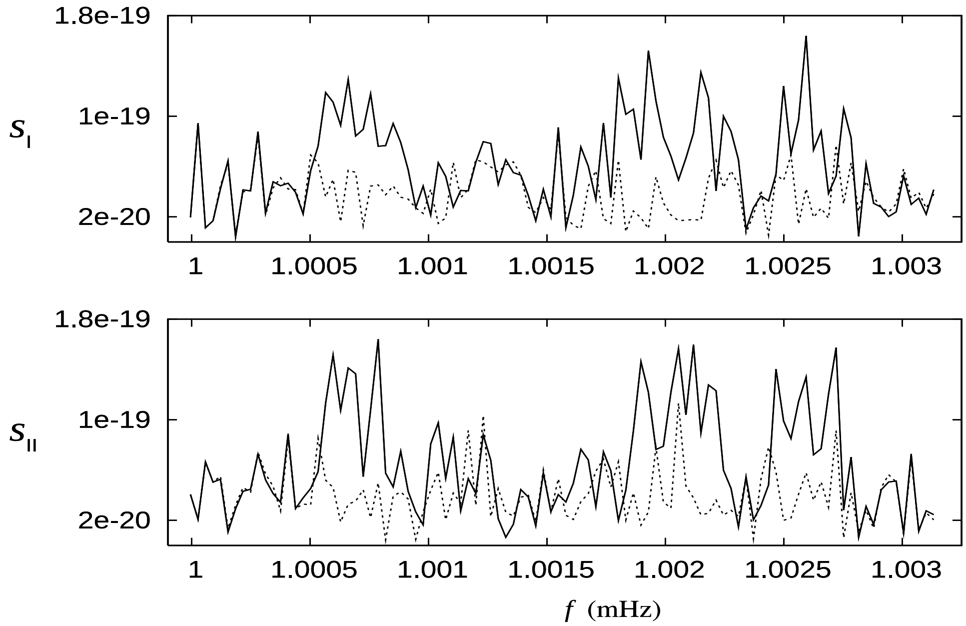

Figure 7.5: Simulated LISA data with 10 galactic binaries. The solid lines show the total signal plus noise, while the dashed lines show the instrument noise contribution.

The current algorithm is intended to give a proof of principle, and is certainly far from optimal. This choice of proposal mixtures was based on a few hundred runs using several different mixtures. There is little doubt that a better algorithm could be constructed that uses a larger variety of proposal distributions in a more optimal mixture.

The improved F-MCMC algorithm was tested on a variety of simulated data sets that included up to 10 sources in a 100 bin snippet (once again one year of observation is being used). The algorithm performed very well, and was able to accurately recover all sources with SNR > 5 so long as the degree of source correlation was not too large. Generally the algorithm could de-blend sources that had correlation coefficients $C_{12} = (h_1|h_2)/\sqrt{(h_1|h_1)(h_2|h_2)}$ below 0.3. A full investigation of the de-blending of highly correlated sources is deferred to a subsequent study. For now one representative example from the 10 source searches is presented.

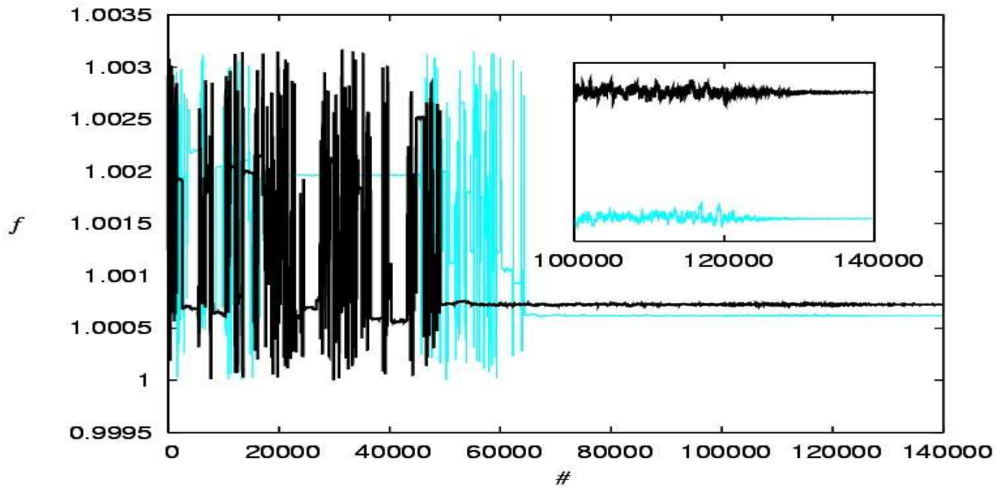

Figure 7.6: Trace plots of the frequencies for two of the model sources. During the annealing phase ($\# < 10^5$) the chain explores large regions of parameter space. The inset shows a zoomed in view of the chain during the MCMC run at $\beta = 1/2$ and the final super cooling which starts at $\# = 1.2 \times 10^5$.

Table 7.3: F-MCMC search for 10 galactic binaries using a model with 10 sources. The frequencies are quoted relative to 1 mHz as $f = 1\,\mathrm{mHz} + \delta f$ with δf in μHz.

	SNR	A (10^{-22})	δf	θ	ϕ	ψ	ι	φ_0
True	8.1	0.56	0.623	1.18	4.15	2.24	2.31	1.45
MCMC ML	8.6	0.76	0.619	1.13	4.16	1.86	2.07	1.01
True	9.0	0.47	0.725	0.80	0.69	0.18	0.21	2.90
MCMC ML	11.3	0.97	0.725	0.67	0.70	0.82	0.99	1.41
True	5.1	0.46	0.907	2.35	0.86	0.01	2.09	2.15
MCMC ML	6.0	0.67	0.910	2.07	0.61	3.13	1.88	2.28
True	8.3	1.05	1.126	1.48	2.91	0.46	1.42	1.67
MCMC ML	6.9	0.75	1.114	1.24	3.01	0.40	1.26	2.88
True	8.2	0.54	1.732	1.45	0.82	1.58	0.79	2.05
MCMC ML	7.7	0.77	1.730	1.99	0.69	1.27	1.18	2.73
True	14.7	1.16	1.969	1.92	0.01	1.04	2.17	5.70
MCMC ML	12.8	1.20	1.964	1.97	6.16	0.97	2.00	6.15
True	4.9	0.41	2.057	2.19	1.12	1.04	2.13	3.95
MCMC ML	5.2	0.66	1.275	0.57	2.81	0.57	1.82	3.93
True	8.8	0.85	2.186	2.21	4.65	3.13	2.01	4.52
MCMC ML	10.0	1.01	2.182	2.43	5.06	0.26	2.00	5.54
True	7.6	0.58	2.530	2.57	0.01	0.06	0.86	0.50
MCMC ML	6.7	0.98	2.582	2.55	6.03	2.71	1.52	5.58
True	11.7	0.69	2.632	1.17	3.14	0.45	2.53	0.69
MCMC ML	13.5	1.39	2.627	1.55	3.07	3.08	1.94	6.07

A set of 10 galactic sources was randomly selected from the frequency range $f \in [0.999995, 1.003164]$ mHz and their signals were processed through a model of the LISA instrument response. The root spectral densities in the two independent LISA data channels are shown in Figure 7.5, and the source parameters are listed in Table 7.3. Note that one of the sources had a SNR below 5. The data was then search using the improved F-MCMC algorithm using a model with 10 sources (70 parameters). The annealing time was set at 10^5 steps, and this was followed by a short MCMC run of 2×10^4 steps and a super cooling phase that lasted 2×10^4 steps. The main MCMC run was kept short, since this portion of the work was focused on extracting maximum likelihood estimates. Figure 7.6 shows a trace plot of the chain that focuses on the frequencies of two of the model sources. During the early hot phase the chain moves all over parameter space, but as the system cools to $\beta = 1/2$ the chain settles down and locks onto the sources. During the final super cooling phase the movement of the chain is exponentially damped as the model is trapped at a mode of shrinking width and increasing height.

The list of recovered sources can be found in Table 7.3. The low SNR source (SNR $= 4.9$) was not recovered, but because the model was asked to find 10 sources it instead dug up a spurious source with SNR $= 5.2$. With two exceptions, the intrinsic parameters for the other 9 sources were recovered to within 3σ of the true parameters (using the FIM estimate of the parameter recovery errors). The two misses were the frequency of the source at $f = 1.00253$ mHz (out by 19σ) and the co-latitude of the source at $f = 1.002632$ mHz (out by 6σ). It is no co-incidence that these misses occurred for the two most highly correlated sources ($C_{9,10} = -0.23$). The full source cross-correlation matrix is listed in (7.4).

$$C_{ij} = \frac{(h_i|h_j)}{\sqrt{(h_i|h_i)(h_j|h_j)}}$$

$$C_{ij} = \begin{pmatrix} 1 & 0.08 & 0 & 0.01 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.08 & 1 & 0.02 & 0.01 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.02 & 1 & -0.06 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.01 & 0.01 & -0.06 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0.01 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -0.03 & 0.03 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -0.03 & 1 & -0.05 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.01 & 0.03 & -0.05 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -0.23 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.23 & 1 \end{pmatrix} \tag{7.4}$$

The MCMC derived maximum likelihood estimates for the the source parameters can be used to regress the sources from the data streams. Figure 7.7 compares the residual signal to the instrument noise. The total residual power is below the instrument noise level as some of the noise has been incorporated into the recovered signals.

Model Selection

In the preceding examples models were used that had the same number of components as there were sources in the data snippet. This luxury will not be available with the real LISA data. A realistic data analysis procedure will have to explore model space as well as parameter space. It is possible to generalize the MCMC approach to simultaneously explore both spaces by incorporating trans-dimensional moves in the proposal distributions. In other words, proposals that change the number of sources

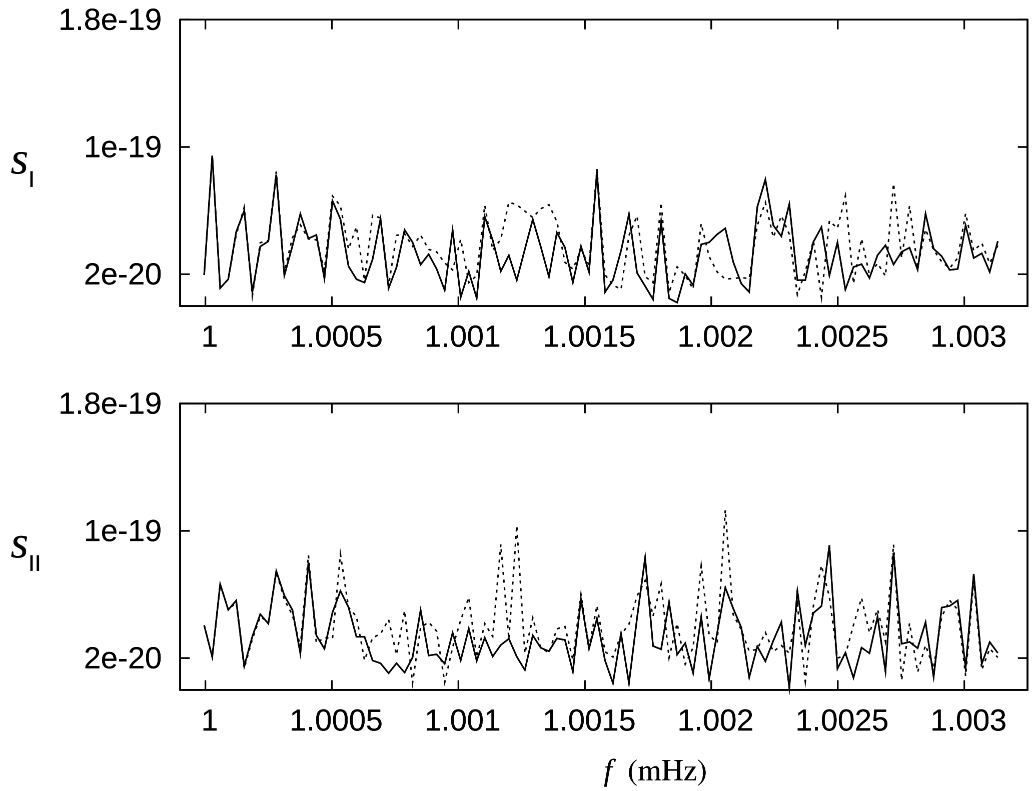

Figure 7.7: The LISA data channels with the sources regressed using the maximum likelihood parameter estimates from the F-MCMC search. The solid lines show the residuals, while the dashed lines show the instrument noise contribution.

being used in the fit. One popular method for doing this is Reverse Jump MCMC [94], but there are other simpler methods that can be used. When trans-dimensional moves are built into the MCMC algorithm the odds ratio for the competing models is given by the fraction of the time that the chain spends exploring each model. While trans-dimensional searches provide an elegant solution to the model determination problem in principle, they can perform very poorly in practice as the chain is often reluctant to accept a trans-dimensional move.

A simpler alternative is to compare the outputs of MCMC runs using models of fixed dimension. The odds ratio can then be calculated using Bayes factors. Calculating the marginal likelihood of a model is generally very difficult as it involves an integral over all of parameter space:

$$p_X(s) = \int p(s|\vec{\lambda}, X) p(\vec{\lambda}, X) d\vec{\lambda} \,. \tag{7.5}$$

Unfortunately, this integrand is not weighted by the posterior distribution, so one cannot use the output of the MCMC algorithm to compute the integral. When the likelihood distribution has a single dominant mode, the integrand can be approximated using the Laplace approximation:

$$p(\vec{\lambda}, X) p(s|\vec{\lambda}, X) \simeq p(\vec{\lambda}_{\rm ML}, X) p(s|\vec{\lambda}_{\rm ML}, X) \exp\left(-\frac{(\vec{\lambda} - \vec{\lambda}_{\rm ML}) \cdot F \cdot (\vec{\lambda} - \vec{\lambda}_{\rm ML})}{2} \right) \tag{7.6}$$

where F is given by the Hessian

$$F_{ij} = \frac{\partial^2 \ln(p(\vec{\lambda}, X) p(s|\vec{\lambda}, X))}{\partial \lambda_i \partial \lambda_j} \Big|_{\vec{\lambda} = \vec{\lambda}_{\rm ML}} \,. \tag{7.7}$$

When the priors $p(\vec{\lambda}, X)$ are uniform or at least slowly varying at maximum likelihood, F_{ij} is equal to the FIM Γ_{ij}. The integral is now straightforward and yields

$$p_X(s) \simeq p(\vec{\lambda}_{\rm ML}, X) p(s|\vec{\lambda}_{\rm ML}, X) \frac{(2\pi)^{D/2}}{\sqrt{\det F}} \,. \tag{7.8}$$

With uniform priors $p(\vec{\lambda}_{\rm ML}, X) = 1/V$, where V is the volume of parameter space, and $(2\pi)^{D/2}/\det F = \Delta V$, where ΔV is the volume of the error ellipsoid. Recalling (3.33) one can see how the marginal likelihood is a product of the maximum likelihood and the Occam factor.

To illustrate how the Bayes factor can be used in model selection, the F-MCMC search described in the previous section was repeated, but this time using a model with 9 sources. The results of a typical run are presented in Table 7.4. The parameters

Table 7.4: F-MCMC search for 10 galactic binaries using a model with 9 sources. The frequencies are quoted relative to 1 mHz as $f = 1\,\mathrm{mHz} + \delta f$ with δf in μHz.

	SNR	A (10^{-22})	δf	θ	ϕ	ψ	ι	φ_0
True	8.1	0.56	0.623	1.18	4.15	2.24	2.31	1.45
MCMC ML	8.6	0.77	0.619	1.12	4.16	1.86	2.06	1.02
True	9.0	0.47	0.725	0.80	0.69	0.18	0.21	2.90
MCMC ML	11.3	0.95	0.725	0.67	0.70	0.84	0.98	1.30
True	5.1	0.46	0.907	2.35	0.86	0.01	2.09	2.15
MCMC ML	6.2	0.75	0.910	2.09	0.61	3.09	1.82	2.21
True	8.3	1.05	1.126	1.48	2.91	0.46	1.42	1.67
MCMC ML	7.0	0.81	1.112	1.24	2.95	0.45	1.31	2.55
True	8.2	0.54	1.732	1.45	0.82	1.58	0.79	2.05
MCMC ML	7.5	0.76	1.730	1.95	0.70	1.23	1.19	2.68
True	14.7	1.16	1.969	1.92	0.01	1.04	2.17	5.70
MCMC ML	12.9	1.23	1.965	1.97	6.17	0.99	1.99	6.11
True	4.9	0.41	2.057	2.19	1.12	1.04	2.13	3.95
MCMC ML	-	-	-	-	-	-	-	-
True	8.8	0.85	2.186	2.21	4.65	3.13	2.01	4.52
MCMC ML	10.0	1.00	2.182	2.41	5.05	0.23	2.00	5.59
True	7.6	0.58	2.530	2.57	0.01	0.06	0.86	0.50
MCMC ML	5.8	0.72	2.536	0.58	5.70	3.04	1.52	4.64
True	11.7	0.69	2.632	1.17	3.14	0.45	2.53	0.69
MCMC ML	13.2	1.21	2.631	1.41	2.97	0.46	2.02	0.50

of the 9 brightest sources were all recovered to within 3σ of the input values, save for the sky location of the source with frequency $f = 1.00253$ mHz. It appears that confusion with the source at $f = 1.002632$ mHz may have caused the chain to favour a secondary mode like the one seen in Figure 7.4. Using (7.8) to estimate the marginal likelihoods for the 9 and 10 parameter models it was found $\ln p_9(s) = -384.3$ and $\ln p_{10}(s) = -394.9$, which gives an odds ratio of $1 : 4 \times 10^4$ in favour of the 9 parameter model. In contrast, a naive comparison of log likelihoods, $\ln \mathcal{L}_9 = 413.1$ and $\ln \mathcal{L}_{10} = 425.7$ would have favoured the 10 parameter model.

It is also interesting to compare the output of the 10 source MCMC search to the maximum likelihood one gets by starting at the true source parameters then applying the super cooling procedure (in other words, cheat by starting in the neighborhood of the true solution). It was found that $p_{\rm cheat}(s) = -394.5$, and $\ln \mathcal{L}_{\rm cheat} = 421.5$, which show that the MCMC solution, while getting two of the source parameters wrong, provides an equally good fit to the data. In other words, there is *no* data analysis algorithm that can fully deblend the two highly overlapping sources.

Future Plans

The next step is to push the existing algorithm until it breaks. Simulations of the galactic background suggest that bright galactic sources reach a peak density of one source per two 1/year frequency bins [48]. Yet to be attempted is a search involving larger numbers of sources as the current version of the algorithm employs the full $D = 7N$ dimensional FIM in many of the updates, which leads to a large computational overhead. The process of modifying the algorithm so that sources are first grouped into blocks that have strong overlap is under development. Each block is effectively independent of the others. This allows each block to be updated separately, while still taking care of any strongly correlated parameters that might impede mixing of the chain.

It would be interesting to compare the performance of the different methods that have been proposed to solve the LISA cocktail party problem. Do iterative methods like gCLEAN [49] and Slice & Dice [96] or global maximization methods like Maximum Entropy have different strengths and weakness compared to MCMC methods, or do they all fail in the same way as they approach the confusion limit? It may well be that methods that perform better with idealized, stationary, Gaussian instrument noise will not prove to be the best when faced with real instrumental noise.

BIBLIOGRAPHY

[1] C. W. Misner, K. S. Thorne, and J. A. Wheeler. *Gravitation*. W. H. Freeman and Company, 1973.

[2] J. C. Maxwell. *Phil. Trans. Roy. Soc. Lon.*, 155:459, 1865.

[3] A. A. Michelson and E. W. Morley. *The American of Science*, 34:333, 1887.

[4] A. Einstein. *Annalen der Physik*, 17:891, 1905.

[5] A. Einstein. *Jahrbuch der Radioaktivitaet und Elektronik*, 4:411, 1907.

[6] A. Einstein. *Sitzungsbber Preuss Akad. Wiss. Berlin*, page 778, 1915.

[7] B.F. Schutz. *gr-qc/0111095*, 2001.

[8] P. L. Bender et al. *LISA Pre-Phase A Report*. 1998.

[9] J. Weber. *Phys. Rev.*, 117:306, 1960.

[10] J. Weber. *Phys. Rev. Lett.*, 17:1228, 1966.

[11] R. Weiss. *Quart. Prog. Rep. RLE MIT*, 105:54, 1971.

[12] G.E. Moss, L. R. Miller, and R. L. Forward. *Appl. Opt.*, 10:2495, 1971.

[13] S. A. Hughes, S Marka, P. L. Bender, and C. J. Hogan. *astro-ph/0110349*, 2001.

[14] N. J. Cornish and S. L. Larson. *Class. Quant. Grav.*, 18:3473, 2001.

[15] M. Tinto, J. W. Armstrong, and F. B. Estabrook. *Phys. Rev. D*, 63:021101(R), 2001.

[16] B.F. Schutz. *Class. Quant. Grav.*, 6:1761, 1989.

[17] B.F. Schutz. *gr-qc/9710080*, 1997.

[18] L. S. Finn. *gr-qc/9903107*, 1999.

[19] P. Jaranowski and A. Krolak. *gr-qc/0204090*, 2002.

[20] X Siemens, J.D.E. Creighton, I. Maor, S. R. Majunder, K. Cannon, and J. Read. *gr-qc/0603111*, 2006.

[21] T. Damour and A. Vilenkin. *Phys. Rev. D*, 64:064008, 2001.

[22] B. Allen and A. C. Ottewill. *Phys. Rev. D*, 63:063507, 2001.

[23] J. Crowder and N. J. Cornish. *Phys. Rev. D*, 72:083005, 2005.

[24] P. L. Bender, P. J. Armitage, M. C. Begelman, and R. Pema. *Massive Black Hole Formation and Growth White Paper submitted to NASA SEU Roadmap Committee*. 2005.

[25] E. S. Phinney et al. *The Big Bang Observer: Direct Detection of Gravitational WavesFrom the Birth of the Universe to the Present, NASA Mission Concept Study*. 2004.

[26] N. J. Cornish, E. S. Phinney, and N. Seto. *in preparation*.

[27] K. S. Thorne. *300 Years of Gravitation*. Cambridge University Press, Cambridge, England, 1987.

[28] B. F. Schutz. *The Detection of Gravitational Waves*. Cambridge University Press, Cambridge, England, 1991.

[29] B.S. Sathyaprakash and S.V. Dhurandhar. *Phys. Rev. D*, 44:3819, 1991.

[30] S.V. Dhurandhar and B.S. Sathyaprakash. *Phys. Rev. D*, 49:1707, 1994.

[31] C. Cutler and É. É. Flanagan. *Phys. Rev. D*, 49:2658, 1994.

[32] R. Balasubramanian and S.V. Dhurandhar. *Phys. Rev. D*, 50:6080, 1994.

[33] B.S. Sathyaprakash. *Phys. Rev. D*, 50:7111, 1994.

[34] T. A. Apostolatos. *Phys. Rev. D*, 52:605, 1996.

[35] E. Poisson and C. M. Will. *Phys. Rev. D*, 52:848, 1995.

[36] R. Balasubramanian, B.S. Sathyaprakash, and S.V. Dhurandhar. *Phys. Rev. D*, 53:3033, 1996.

[37] B. J. Owen. *Phys. Rev. D*, 53:6749, 1996.

[38] B. J. Owen and B.S. Sathyaprakash. *Phys. Rev. D*, 60:022002, 1999.

[39] L. S. Finn. *Phys. Rev. D*, 46:5236, 1992.

[40] P. Jaranowski and A. Krolak. *Phys. Rev. D*, 49:1723, 1994.

[41] P. Jaranowski, A. Krolak, and B. F. Schutz. *Phys. Rev. D*, 58:063001, 1998.

[42] M. Tinto and J. W. Armstrong. *Phys. Rev. D*, 59:102003, 1999.

[43] L. J. Rubbo, N. J. Cornish, and O. Poujade. *Phys. Rev. D*, 69:082003, 2004.

[44] C. Cutler. *Phys. Rev. D*, 57:07089, 1998.

[45] T. A. Prince, M. Tinto, S. L. Larson, and J. W. Armstrong. *Phys. Rev. D*, 66:122002, 2002.

[46] C. W. Helstrom. *Statistical Theory of Signal Detection*. London, 2nd edition edition, 1968.

[47] L.A. Wainstein and V.D. Zubakov. *Extraction of Signals from Noise*. Prentice-Hall, Englewood Cliffs, 1962.

[48] S. Timpano, L. J. Rubbo, and N. J. Cornish. *gr-qc/0504071*, 2005.

[49] N. J. Cornish and S. L. Larson. *Phys. Rev. D*, 67:103001, 2003.

[50] J. Högbom. *Ap.J Suppl.*, 15:417, 1974.

[51] N. J. Cornish. Talk given at GR17, Dublin, July, 2004.

[52] W. R. Gilks. *Markov Chain Monte Carlo in Practice*. Chapman and Hall, London, 1996.

[53] D. Gamerman. Chapman and Hall, London, 1997.

[54] N. J. Cornish and J. Crowder. *Phys. Rev. D*, 72:043005, 2005.

[55] J. Holland. *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, Michigan, 1975.

[56] J. Crowder, N. J. Cornish, and J. L. Reddinger. *Phys. Rev. D*, 73:063011, 2006.

[57] M. H. A. Davis. *Gravitational Wave Data Analysis*. Kluwer Academic, Boston, 1989.

[58] G. Schwarz. *Ann. Stats.*, 5:461, 1978.

[59] A. Krolak, M. Tinto, and M. Vallisneri. *Phys. Rev. D*, 70:022003, 2004.

[60] A. J. Farmer and E. S. Phinney. *Mon. Not. Roy. Astron. Soc.*, 346:1197, 2003.

[61] D. N. Spergel et al. *Ap. J. Suppl.*, 148:175, 2003.

[62] L. Blanchet, B. R. Iyer, C. M. Will, and A. G. Wiseman. *Class. Quant. Grav.*, 13:575, 1996.

[63] B. Kocsis, Z. Frei, Z. Haiman, and K. Menou. *astro-ph/0505394*, 2005.

[64] S. A. Hughes and D. E. Holz. *Class. Quant. Grav.*, 20:S65, 2003.

[65] N. Seto. *Phys. Rev. D*, 69:022002, 2004.

[66] J. Crowder and N. J. Cornish. *Phys. Rev. D*, 70:082004, 2004.

[67] D. Hils, P. L. Bender, and R. F. Webbink. *ApJ*, 360:75, 1990.

[68] T. A. Moore and R. W. Hellings. *Phys. Rev. D*, 65:062001, 2002.

[69] R. Takahashi and N. Seto. *ApJ*, 575:1030, 2002.

[70] S. A. Hughes. *Mon. Not. Roy. Astron. Soc.*, 331:805, 2002.

[71] N. Seto. *Phys. Rev. D*, 66:122001, 2002.

[72] A. Vecchio. *Phys. Rev. D*, 67:022001, 2003.

[73] L. Barack and C. Cutler. *Phys. Rev. D*, 69:082005, 2004.

[74] N. J. Cornish. *gr-qc/0304020*, 2003.

[75] E. S. Phinney. *Talk given at the Fourth International LISA Symposium*, July 2002.

[76] R. Hellings. *Comments made at the International Conference on Gravitational Waves: Sources and Detectors*, 1996.

[77] C. E. Shannon. *Bell System Technical Journal*, 27:379, 1948.

[78] J. R. Gair, L. Barack, T. Creighton, C. Cutler, S. L. Larson, E. S. Phinney, and M. Vallisneri. *Class. Quant. Grav.*, 21:S1595, 2004.

[79] S. D. Mohanty and R. K. Nayak. *gr-qc/0512014*, 2005.

[80] D. Hils, P. L. Bender, and R. F. Webbink. *ApJ*, 360:75, 1990.

[81] D. Hils and P. L. Bender. *ApJ*, 537:334, 2000.

[82] C. R. Evans, I. Iben, and L. Smarr. *ApJ*, 323:129, 1987.

[83] V. M. Lipunov, K. A. Postnov, and M. E. Prokhorov. *A&A*, 176:L1, 1987.

[84] G. Nelemans, L. R. Yungelson, and S. F. Portegies Zwart. *A&A*, 375:890, 2001.

[85] L. Barack and C. Cutler. *Phys. Rev. D*, 70:122002, 2004.

[86] C. Darwin. *The Origin of Species*. J. Murray, London, 1859.

[87] N. Christensen and R. Meyer. *Phys. Rev. D*, 58:082001, 1998.

[88] N. Christensen, R. J. Dupuis, G. Woan, and R. Meyer. *Phys. Rev. D*, 70:022001, 2004.

[89] R. Umstatter, N. Christensen, M. Hendry, R. Meyer, V. Simha, J. Veitch, S. Viegland, and G. Woan. *Class. Quant. Grav.*, 22:S901, 2005.

[90] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. *J. Chem. Phys.*, 21:1087, 1953.

[91] W. K. Hastings. *Biometrics*, 57:97, 1970.

[92] C. Andrieu, N. De Freitas, A. Doucet, and M. Jordan. *Machine Learning*, 50:5, 2003.

[93] L. Tierney and A. Mira. *Statistics in Medicine*, 18:2507, 1999.

[94] P. J. Green and A. Mira. *Biometrika*, 88:1035, 2001.

[95] N. J. Cornish and E. K. Porter. *Class. Quant. Grav.*, 22:S927, 2005.

[96] N. J. Cornish, L. J. Rubbo, and R. Hellings. *in preparation*, 2006.

APPENDIX A: LIST OF ACRONYMS

ALIA	Advanced Laser Interferometer Antenna
ALIAS	Advanced Laser Interferometer Antenna in Stero
AURIGA	Antenna Ultracriogenica Risonante per l'Indagine Gravitazionale Astronomica
BBO	Big Bang Observer
BH	Black Hole
BIC	Bayesian Information Criterion
COBE	COsmic Background Explorer
EMRI	Extreme Mass Ratio Inspiral
FIM	Fisher Information Matrix
GA	Genetic Algorithm
GGA	Genetic-Genetic Algorithm
GEO600	Gravitational wave Earth Observatory 600
GMU	Geometric Mean of the parameter Uncertainties
GMUR	Geometric Mean Uncertainty Ratio
GRAIL	Gravitational Radiation Antenna In Leiden
GWB	Gravitational Wave Background
IMBH	Intermediate Mass Black Hole
LIGO	Laser Interferometer Gravitational wave Observatory
LISA	Laser Interferometer Space Antenna
LISAS	Laser Interferometer Space Antenna in Stereo
MCMC	Markov Chain Monte Carlo
MBH	Massive Black Hole
ML	Maximum Likelihood
NS	Neutron Star
PMR	Parameter Mutation Rate
PUR	Parameter Uncertainty Ratio
ROG	Rome One Group
SNR	Signal to Noise Ratio
SGN	Stationary, Gaussian Noise
SMBH	Supermassive Black Hole
TAMA300	Tokyo Advanced Mediumscale Antenna 300
TT	Transverse Traceless
WMAP	Wilkinson Microwave Anisotropy Probe