LISA Data Analysis Status

LISA Mission Science Office

Date: 5 May 2009
Registration: LISA-MSO-TN-1001-2-1
Issue: 2
Revision: 1

The electronic version of this document contains hyperlinks that are indicated by color. Internal links refer to content *within* this document, external links refer to content *outside* this document and require a connection to the internet.

Contribution to this document by:
S. Babak
J. Baker
M. Benacquista
N. Cornish
S. Finn
M. Hewitson
O. Jennrich
A. Królak
S. Larson
M. Poessel
E. Porter
B. Schutz
M. Vallisneri
A. Vecchio
Please direct comments and questions to O. Jennrich.

Change Log

Change	Date	Version	Page
Data rates corrected	5 May 2009	2.1	85

Executive Summary

The analysis of the LISA data spans the entire distance from the raw measurements of the onboard phasemeters to the production of reliable source catalogs.

The reduction of the LISA raw phase measurements to clean Gravitational Wave (GW) observables is well understood, and it has been implemented and tested on the ground. LISA data reduction builds on the legacy of spacecraft Doppler-tracking experiments, which have inspired Time Delay Interferometry (TDI), a technique to efficiently suppress laser frequency noise, exposing all available information about the GW strains. TDI is implemented on the ground as a post-processing step, so no complicated onboard logic is required. No calibration is needed, either, since the TDI-cleaned LISA phasemeter outputs amount to naturally calibrated measurements of the changes in distance between the LISA proof masses. LISA will simultaneously observe all sources all the time, with scheduled interruptions only for short periods of time, as needed for communications and spacecraft maintenance. This eliminates many operational constraints such as time allocations for dedicated observations, and the prioritization of science objectives.

The extraction of the physical and astrophysical source parameters from the GW observables is tackled with a variety of methods, the workhorse of which is the technique of *matched filtering*: the measured data are compared with theoretically modeled waveforms (*templates*), searching parameter space for the best fit. Clearly, extracting reliable source parameters requires very accurate templates. Fortunately, the physics of GW sources (and the corresponding analytical and numerical treatments of Einstein's field equations) have been the subject of intense research over the last thirty years.

As a matter of fact, analytic waveform expansions are known to a very high degree of accuracy for systems such as the low-mass binaries in our Galaxy, or the Massive Black Hole Binary (MBHB)s in the early stage of their evolution; full numerical solutions for the later merger phase of MBHBs have recently become available. Future developments in computational power and algorithmic efficiency will enable highly accurate waveforms for all LISA GW sources, even in their late-stage, strong-gravity regimes.

In addition, the science analysis of LISA data benefits from the tremendous legacy knowledge from the field of time-series analysis, and of course from the thriving analysis program for data from ground-based experiments such as the NSF-funded Laser Interferometer Space Antenna (LIGO) and the European Virgo and GEO600 projects. Compared to the data from the ground-based interferometers, where detected GW signals will be well isolated, either as brief bursts or sharp lines, the LISA data will be rich with a great variety of superimposed, always-on signals, with Signal to Noise Ratio (SNR)s (*i. e.*, intrinsic strengths) from the detection threshold (SNR $\sim$ 5) to more than 10 000. From a measurement viewpoint, such a large dynamical range is easily handled by the LISA phasemeters. From a data-analysis viewpoint, accurate templates (as made possible by the purity of the Einstein field equations) ensure the precise removal of the strong foreground signals; and efficient global-fitting methods (such as Markov Chain Monte Carlo algorithms) can deal with all resolvable signals simultaneously.

LISA data analysis is already being practiced (on synthetic data including noise and numerous sources) in the ongoing Mock LISA Data Challenge (MLDC), which joins more than ten research groups around the world. This program includes data-analysis challenges of increasing complexity and reality, leading to the final goal of analyzing a realistic representation of a complete LISA data set.

LISA data analysis will be conducted jointly by ESA and NASA, with details subject to a future MoU between the agencies. However, we already know that the complete LISA data set, including science data and auxiliary data, will be about 1 TByte for a five-year mission. Thus, it will be possible to release to the general public the full science data, including satellite ephemerides and a catalog of identified sources and their parameters, as well as software tools and algorithms, according to the rules governing the two agencies.

Contents

1. General

1.1. Mission description

LISA is a mission to detect and observe low-frequency gravitational waves. It comprises three identical spacecraft whose positions mark the vertices of an equilateral triangle five million km on a side, in an Earth-like orbit around the Sun. LISA performs interferometric measurements of the distances between proof masses that are in a purely gravitational orbit, shielded from solar pressure and other forces by the spacecraft. The triangular structure of the constellation allows one to perform three independent distance measurement that, suitably combined, enable LISA to simultaneously measure both polarizations of a gravitational wave (GW).

The center of the LISA triangle traces an orbit in the ecliptic plane, 1 AU from the Sun and trailing the Earth, with the plane of the triangle inclined at 60° to the ecliptic. The orbits of the spacecraft are arranged so that the constellation is maintained throughout the year without the need of additional corrective maneuvers; the triangle counter-rotates about its center once per year.

The measurement principle of LISA is based on optical interferometry. Each spacecraft receives light from each of the two other spacecraft and combines it with a local oscillator, registering the phase of the incoming light. As gravitational waves cause a time-varying *strain* in space-time, they change the proper distance between free-falling masses and hence the phase of light traveling between these masses. For light paths short compared to the gravitational wavelength, this change is proportional to the absolute distance between the proof masses, which calls for a baseline as long as possible. However, if the gravitational wave period is short enough that the light interacts with the wave for one full cycle of its period, the light will experience no net phase shift, as the shortening and lengthening of the optical path length will compensate.

Therefore, the arm length of LISA has been chosen to provide a measurement band from 3×10^{-5} Hz to 0.1 Hz. In this frequency range falls an abundance of sources such as coalescing massive black holes, among the strongest sources LISA will detect with an SNR of many thousand, large numbers ($> 10\,000$) of compact binary systems in our galaxy, as well as hundreds of compact objects falling into massive black holes at the center of galaxies, all with SNR above the detection threshold (SNR > 5).

All three spacecraft can be launched together by a single ATLAS V (531). Each spacecraft carries small, steerable antennas used for transmitting the science and engineering data, stored on board for up to six days, in the Ka-band (31.8 GHz to 32.3 GHz) to the NASA Deep Space Network (DSN). The nominal mission lifetime is five years.

1.2. LISA data analysis overview

LISA, like other gravitational-wave detectors, observes gravitational waves *coherently*, which means that it follows the phase of the wave. Indeed, because gravitational waves are generated by the large-scale mass motions in their sources, the phase normally contains most of the information of interest: parameters of the source and its evolution. LISA consists of three differently oriented quadrupolar detectors, so it has a nearly isotropic sensitivity over the whole sky. It separates signals from different directions by using the information from the phase and amplitude modulation induced by its motion around the Sun. LISA has in addition enormous dynamic range, at least a factor of 10^5 in amplitude (or 10^{10} in energy flux).

The main techniques employed in data analysis for coherent signals embedded in noise is the *matched filter*, where a template representing the signal with its expected phase evolution is correlated with the data. If the match is statistically significant and other systematics are under control then a detection can be claimed. This is the way that the ground-based gravitational wave detector projects (LIGO, GEO, VIRGO) are currently searching for signals whose raw amplitudes are well below their detectors' noise output.

Normally, the templates of expected signals belong to families that are characterized by a certain set of parameters. Internal parameters are those that determine the emitted waveform of the source (*i. e.*, the physics of the source), such as masses, spins and orbital parameters of the constituents. External parameters are those that determine the signal caused by the received waveform in the detector, such as source position, distance, and orientation on the sky. Members of these families are distinguishable from one another if their phase evolutions over the course of an observation differ by a significant amount. By finding the best match, data analysis automatically measures the best values of the parameters of the template family.

For ground-based detectors the signals are so weak that they do not significantly interfere with one another: the problem of detecting each signal is to find it against instrumental noise, no matter how many other signals may be present. Therefore, the ground-based data analysis efforts involve many parallel searches for different families of signals. For LISA, the instrumental sensitivity is so high (and noise so low) that signals must be recognized against a background of other signals. This added complication means that all families must be searched for in a unified way, since the detection of a signal in one family can affect the search for signals in other families. Clearly, LISA benefits from some legacy from the development of ground-based data analysis methods, but in addition it has to develop its own methods for performing global searches for signals.

Template-based searches require a good theoretical understanding of the emitting sources. For most of LISA's expected sources, the waveforms are known today with exquisite precision. Waveforms for binary systems (such as the low-mass binaries in our galaxy or the massive–black-hole binaries in the early stages of their evolution) are known in the PN approximation up to orders of $(v/c)^7$. This is sufficient for detecting these sources in LISA's frequency band. Black-hole ringdown radiation is also very well modelled. Rapid improvements in numerical relativity in the last couple of years make it very likely that the very interesting merger phase of black holes will be modelled well enough for LISA to use observations to test strong-field general relativity. Models for the signals from compact objects being captured by massive black holes are good enough now for detection, and are expected to improve over the next few years to allow precise measurement of parameters.

Of course, not all signals are expected to be theoretically well modelled before LISA flies. A simple example is a stochastic background or foreground signal, either from the early Universe or from a superposition of a large number of distant astronomical sources. This can be detected by LISA if it is stronger than instrumental noise. LISA has a remarkable ability to discriminate instrumental noise from gravitational noise, as is described later in this document.

Even more interesting will be sources that are not modelled beforehand, which comprise the so-called LISA discovery space. LISA's high sensitivity makes it very possible that even signals from very distant unexpected sources might exceed instrumental noise and be detectable. Searching for these will be an important part of LISA's data-analysis activity.

A compact introduction to the concepts employed in LISA Data analysis can be found in Vallisneri (2008).

1.3. Organization of the document

This document is intended to give an overview of the issues in the data analysis for LISA, the present state of the art in dealing with these issues and the plans and expectations for the future.

As many of the issues in LISA data analysis transcend individual source classes, the document contains a description of the *general problems* and their solutions in chapter 2. Those issues that are specific to individual source classes are described in chapters 3 to 6. Each chapter starts with a description of the sources, describes the science return from their detection, explains the characteristics of the signal, the goals and requirements for the data analysis and the recent progress in the area of data analysis. This structure does not follow the LISA Science Objectives as set in the Science Requirement Document (*cf.* Appendix B), but is better suited to describe the data analysis.

The plans for further evolving data-analysis techniques with the help of the scientific community at large are outlined in chapter 8.

The reduction and flow of the data form an integral part of the data analysis for LISA. The techniques involved in this are described in appendix A.

1.4. Configuration Management

Changes to this document shall be controlled using the procedures of the LISA Project Configuration Management Procedure LISA Project office (2001).

1.5. Applicable Documents

Documents directly applicable to this document are

- LISA Science Case Document (LISA Mission Science Office, 2006)
- LISA Science Requirements Document (LISA International Science Team, 2006)
- LISA Mission Requirement Document (LISA Project Office, 2004)

1.6. Acronyms

Please see Appendix C.

1.7. References

LISA International Science Team. Science requirements document, December 2006.

LISA Mission Science Office. LISA: Probing the Universe with Gravitational Waves, December 2006.

LISA Project office. *LISA Configuration Management Procedure*, 2001.

LISA Project Office. LISA Mission Requirement Document, November 2004. LISA-MRD-001.

M. Vallisneri. A LISA Data-Analysis Primer. *ArXiv e-prints*, December 2008, also as arXiv:0812.0751.

2. Global Analysis

2.1. Overview

The LISA Observatory will continuously monitor the entire sky across a frequency range of about 3×10^{-5} Hz to 0.1 Hz. The sensitivity of the instrument is such that many sources can be detected at distances right out to the edge of the observable Universe, and the timescales for signals showing a variation in frequency are such that most signals will remain within the observable frequency band for a period of months or years. Thus, the three independent LISA data streams are expected to contain tens of thousands of signals. The job of the data analyst is to reliably isolate and identify as many of these signals as possible, where identification includes determining both the signal class (for example, inspiralling supermassive black holes) and the signal's location within the parameter space of that class (*e. g.*, the black holes' masses). As such, source-specific issues notwithstanding, a central problem of LISA data processing is how to deal with the multiplicity of simultaneous signals, and this is the topic of the present chapter.

Given the wealth of astrophysical signals, the LISA data-analysis problem is similar to the difficulties encountered in the telecommunications industry when multiple telephone conversations are carried on a single transatlantic cable. The key difference between this situation and LISA data analysis is that we do not get to control the signal encoding. Fortunately, nature itself provides an encoding that is amenable to the identification of distinct sources, as the signals from different sources have very small overlap with one another. For example, if two black-hole–binary systems have parameters that differ by just a few percent, the signal overlap[1] is almost zero. Another way of saying this is that the signals from binary black holes have very rich encodings. The signals with the poorest encoding and the smallest bandwidth stem from the galactic binaries, but these systems will be spread out across the LISA band, and Monte Carlo studies indicate that less than 1% of the bright galactic signals will overlap with one another with greater than 50% correlation (Crowder and Cornish, 2006).

It is a key feature of the LISA data that the measurements are phase coherent. The most important consequence of this is that, in principle, the presence of bright signals does not preclude the reliable extraction of weaker signals. In this regard,

[1] Here, the term *overlap* is used in the conventional sense from gravitational wave data analysis, *i. e.* the product of the two signals integrated over the observation time: $(s_1|s_2) = \frac{\int_0^T s_1(t)s_2(t)dt}{\int_0^T s_1^2 dt \int_0^T s_2^2 dt}$.

LISA data analysis differs from the techniques employed in optical astronomy, where powers add, not amplitudes, and in contrast shares key traits with radio astronomy. While phase coherence allows in principle for lossless subtraction, this can only be achieved in practice using sophisticated analysis algorithms. More specifically, one needs to have access to accurate theoretical waveform templates for each signal type, and one must perform a coherent global fit using a "multi-template" that combines in a single entity the templates for each component of the model.

For the LISA multi-template, it is anticipated that the total number of parameters will be of the order 100 000. The great majority of those correspond to between 10 000 and 30 000 resolvable galactic binary systems, each described by 7, 8 or 9 parameters. While these numbers may seem daunting, they are comparable to the model dimensions encountered in areas such as financial-market modeling or geophysics. Moreover, the full global solution can be approached in a hierarchical or iterative manner, possibly utilizing sub-optimal methods, which do not rely on coherent matched filtering.

The LISA global analysis problem belongs to a category of problems known in the statistics community as "mixture models". In a mixture model, one not only has to determine the parameters that describe the components of the mixture, but also the number of resolvable components making up the mixture. Models with more components will generally produce a better fit to the data, but these additional components might just provide a fit to certain features of the instrument noise. The goal is to find the most parsimonious model that strikes a balance between model complexity and quality of fit.

In the past few years considerable progress has been made on global data analysis techniques for LISA, and several promising approaches have been developed and tested on simulated LISA data. Full-scale algorithm development and testing is now in progress under the auspices of the Mock LISA Data Challenges (*cf.* chapter 8.2).

2.2. Global Analysis Issues

The output of each LISA data stream can be written as

$$s_\alpha(t) = h_\alpha(t, \vec{\lambda}) + n_\alpha(t) = \sum_{i=1}^{N} h_\alpha^i(t, \vec{\lambda}_i) + n_\alpha(t)\,. \tag{2.1}$$

Here $h_\alpha^i(t, \vec{\lambda}_i)$ describes the response registered in detector channel α to a source with parameters $\vec{\lambda}_i$. The quantity $h_\alpha(t, \vec{\lambda})$ denotes the combined response to a collection of N sources with total parameter vector $\vec{\lambda} = \cup_i \vec{\lambda}_i$ and $n_\alpha(t)$ denotes the instrument noise in channel α. Extracting the parameters of each individual source from the combined response to all sources defines the LISA global analysis problem.

In practice it will be impossible to resolve all of the millions of signals that contribute to the LISA data streams and indeed there is no requirement to do so. 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 is to provide estimates for the parameters describing each of the M sources that have integrated signal-to-noise ratios above some threshold, where it is understood that the noise includes the instrument noise, residuals from the regression of bright sources, and the signals from unresolved sources.

The optimal filter for the LISA signals is a multi-template describing all M resolvable sources. The number of parameters required to describe an individual source ranges from seven for a slowly evolving circular galactic binary to seventeen for eccentric, spinning black-hole binaries. The number of resolvable sources is estimated to be around 10^4, so the full parameter space has dimension $D \sim 10^5$.

The key features of global data analysis for LISA are:

1. The large dimension of the parameter space;
2. The non-orthogonality of the individual signals;
3. The number of resolvable signals, which is *a priori* unknown;
4. The need for accurate theoretical templates for each component of the model and
5. The complexity of some of the signals.

Before we can undertake the construction of the global template, it is necessary to know its building blocks—in other words, to develop techniques for detecting and characterizing each of the different types of signal that LISA can expect to measure. Thus, the first step of developing LISA data-analysis methods is to develop source-specific analysis techniques, which are described in other sections of this document (chapters 3 to 6). Only as a second step will they be integrated into a global analysis pipeline. Note that the complexity of some of the individual signals, such as EMRIs (*cf.* section 4), may necessitate intermediate steps in the analysis, such as methods that are mathematically sub-optimal but computationally less demanding for the initial detection, followed by the application of a more computationally expensive matched filtering step after a first rough estimate of the parameters has been made. One of the major areas of development will concern the creation of a global data-analysis pipeline in which these hierarchical searches are embedded in the overall global analysis framework. Note that, in one regard, the complexity of those signals which are at present less well understood will actually work in our favor: While errors in the theoretical templates will lead to residuals that can distort the global fit, these complex signals, and hence their residuals, have very little overlap with other signals in the data stream.

Indeed, the complexity and redundancy (think multiple harmonics) in the encoding of the EMRI and BBH signals make them relatively easy to isolate in the global fit. The main challenge to global data analysis comes from the galactic foreground, where we are faced with a large number of low-bandwidth signals. On the other hand, the templates for individual galactic binaries are well understood, computationally inexpensive to generate and easy to search for. And while the signal encoding is relatively poor, Monte Carlo studies of simulated galactic foregrounds show that typical bright signals (SNR > 10) have small overlaps with one another: 90% have overlaps of less than 10% with any other bright galactic binary.

2.3. Global Analysis Techniques

To date, most studies of LISA global analysis issues have focused on the galactic foreground, though there have been some studies that have looked at combining BBH and galactic binary searches. The various approaches can broadly be categorized as iterative or simultaneous, though some approaches combine elements of both.

Examples of iterative methods are the "*gCLEAN*" and "*Slice & Dice*" algorithms; the former is based on the CLEAN (Högbom, 1974) algorithm used in radio astronomy. The *gCLEAN* (Cornish and Larson, 2003) algorithm identifies the brightest source in the data stream and uses a template to subtract a small amount of this signal. The procedure is iterated until some stopping criteria are met, and the sources are reconstructed from the list of partial subtractions. The *Slice & Dice* (Rubbo et al., 2006) algorithm identifies and fully subtracts the brightest source in the data stream (the Slice step). At the next iteration the second brightest signal is subtracted, then a non-linear least squares procedure is used to simultaneously re-fit both signals (the Dice step). The procedure is iterated until some stopping criteria are met, with a simultaneous re-fit to all the candidate sources being performed at each iteration. The re-fitting is designed to correct for errors introduced at the Slice step where possible correlations between signals are ignored. The identification of bright sources is generally achieved using a hierarchical template grid.

A possible drawback to the iterative approaches is that highly correlated signals might send the solution off in the wrong direction. Another issue for the iterative methods is the choice of stopping criteria. These concerns have lead to the the development of simultaneous-analysis techniques belonging to the Markov Chain Monte Carlo (MCMC) family, such as Metropolis-Hastings sampling (Christensen et al., 2004; Hastings, 1970; Metropolis et al., 1953) and Reverse Jump Markov Chain Monte Carlo (Wickham et al., 2006).

A related approach is provided by Genetic Algorithms (Crowder et al., 2006). This method implements the idea of "survival of the fittest" from Darwin's theory of evolution by natural selection. The set of parameters of many signals is modeled as an *evolving organism* and its *fitness* is measured by the $\mathcal{F}$-statistic.

Other simultaneous analysis techniques include Maximum Entropy (Larson and Finn, 2006) and Tomographic methods (Mohanty and Nayak, 2006).

The philosophy behind the MCMC techniques is that they alleviate the "curse of dimensionality" by providing search techniques that scale linearly, rather than geometrically, with the search dimension. A multi-template is constructed that describes some large number of sources, and the source parameters are updated before the next iteration using either a Metropolis-Hastings step or Darwinian selection. The number of sources can be increased or decreased at each step (in other words, the model dimension becomes a search parameter), and quantities that describe the noise in each channel (such as the spectrum and cross-channel correlation) can also be included as search parameters. The end result is an algorithm that can simultaneously estimate the number of resolvable signals, characterize the instrument noise, recover the source parameters and provide error estimates on the parameters. The MCMC approach has become the standard analysis method in many different fields of study that share elements in common with the global LISA data analysis, and it is likely to play an important role in the future development of LISA data analysis.

2.4. References

N. Christensen, R. J. Dupuis, G. Woan, and R. Meyer. Metropolis-Hastings algorithm for extracting periodic gravitational wave signals from laser interferometric detector data. *Phys. Rev. D*, 70(2):022001–+, July 2004, also as arXiv:gr-qc/0402038.

N. J. Cornish and S. L. Larson. LISA data analysis: Source identification and subtraction. *Phys. Rev. D*, 67:103001–+, May 15 2003, also as arXiv:astro-ph/0301548.

J. Crowder and N. Cornish. A Solution to the Galactic Foreground Problem for LISA. *ArXiv Astrophysics e-prints*, November 2006, also as arXiv:astro-ph/0611546.

J. Crowder, N. J. Cornish, and J. L. Reddinger. LISA data analysis using genetic algorithms. *Phys. Rev. D*, 73:063011–+, March 2006, also as arXiv:gr-qc/0601036.

W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.

J. A. Högbom. Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines. *A&AS*, 15:417–+, June 1974.

S. L. Larson and L. S. Finn. The resolving power of LISA: comparing techniques for binary analysis. In S. M. Merkowitz and J. C. Livas, editors, *LASER INTER-*

FEROMETER SPACE ANTENNA: 6th International LISA Symposium, volume 873, pages 415–421. AIP, 2006.

N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953.

S. D. Mohanty and R. K. Nayak. Tomographic approach to resolving the distribution of LISA Galactic binaries. *Phys. Rev. D*, 73:083006–+, April 2006, also as arXiv:gr-qc/0512014.

L. J. Rubbo, N. J. Cornish, and R. W. Hellings. Slice & Dice: Identifying and Removing Bright Galactic Binaries from LISA Data. In S. M. Merkowitz and J. C. Livas, editors, *LASER INTERFEROMETER SPACE ANTENNA: 6th International LISA Symposium*, volume 873, pages 489–493. AIP, August 2006, also as arXiv:gr-qc/0608112.

E. D. L. Wickham, A. Stroeer, and A. Vecchio. A Markov chain Monte Carlo approach to the study of massive black hole binary systems with LISA. *Class. Quantum Grav.*, 23:819–+, October 2006.

3. Massive Black Hole Binaries (MBHB)

3.1. Description of the sources

The observations of the sub-pc central regions of the external galaxies conducted with space-borne observatories and from the ground using adaptive optics indicate presence of the massive ($10^6\, M_\odot \ldots 10^9\, M_\odot$) dark compact objects, which are most probably massive black holes (MBH), in the center of most galaxies. The hierarchical structure formation model, which is currently most favorable, suggests that todays MBH, are the direct result of the coalescence of smaller "seed" black holes of about $10^4\, M_\odot \ldots 10^5\, M_\odot$ throughout cosmic history (Hopkins et al., 2006). This implies that coalescence of the MBHs should be relatively frequent events especially at high redshift.

The inspiral of two intermediate mass black holes that are at the heigh end of their mass range ($10^4\, M_\odot \ldots 10^5\, M_\odot$) will be also detectable by LISA to quite high distances, allowing to etsablish the existence of those objects or set a stringent upper bound on the rates.

Direct observation of massive black hole binaries in the electromagnetic spectrum have become more numerous in the recent past. The direct observation of the radio galaxy 0402+379 suggests that at its center there are two MBHs with a total mass $10^8\, M_\odot$ and an orbital separation of about 7 pc (Rodriguez et al., 2006). The DEEP2 Galaxy redshift survey observed two MBH with masses of a about $5 \times 10^5\, M_\odot$ and $5 \times 10^6\, M_\odot$, respectively, with a seperation of about 1.2 kpc in EGSD2 J142033.66+525917.5 at a redshift of $z = 0.709$. Quite recently SDSS J153636.22+044127.0 has been discovered to host two MBH with a seperation of 0 1pc and masses of $7.9 \times 10^8\, M_\odot$ and $2 \times 10^7\, M_\odot$, respectively. This object is particularly interestin as it has a comparatively low redshift ($z \sim 0.38$) (Boroson and Lauer, 2009).

The signals emitted by massive black hole binaries (MBHB) are strong and evolve throughout the whole LISA's frequency band as holes approach each other and merge. The MBHs in the binaries are expected to be spinning so spin-spin and spin-orbital interaction adds complexity to the signal. The binary systems with unequal masses (especially those with mass ratio reaching 100) might have a measurable orbital eccentricity as they enter LISA's band. The final product of the merger of two MBHs is a MBH that spins rapidly, *i. e.* possesses dimensionless spin $a > 0.5$.

3.2. Science return from detecting these sources

The observations of MBHB coalescences will address several of LISA's science objectives. Firstly, detection of the MBHB signals themselves will provide direct observations of the black holes. Measuring the spins and masses of the MBHs will give us valuable information about the mechanism of their formation: rapid spins will imply that much of the MBHs mass were built up by gas accretion from a disk, moderate spins imply building the MBHs by a sequence of major mergers of comparable mass MBHs and the low spins imply that MBHs are mostly built by capturing smaller objects coming from random directions. Since we will be able to detect the merger of MBHB out to high redshift ($z \geqslant 10$), we should be able to test the viewpoint that modern galaxies resulted from the merger of smaller "seed" galaxies in the early Universe. It seems hard to get the same information from any other route. Knowing the parameter values of the central objects will also enable more accurate studies of the dynamics of the stellar populations in the bulge (Haehnelt, 2003; Portegies Zwart et al., 2006). Furthermore, we will be able to make very accurate distance measurements which enables us to conduct high-precision cosmology by using MBHB signals as standard sirens[1] (Dalal et al., 2006; Deffayet and Menou, 2007; Holz and Hughes, 2005; Kocsis et al., 2006; Linder, 2008).

MBHBs will serve as laboratories for fundamental tests of gravitational theory. The measurement of their masses and spins will confirm (or disprove) some of the untested predictions of GR (Arun et al., 2006; Berti et al., 2005a,b, 2006; Poisson, 1996). We should be able to probe predictions of General Relativity in the different stages of binary evolution starting with a moderately relativistic inspiral phase to a nonlinear strong-field regime. From analyzing the dispersion of the inspiral gravitational wave signal we can draw conclusion about a non-zero graviton mass (Will and Yunes, 2004). Detecting and characterizing the post-merger phase, where the resulting black hole sheds irregularities and deformations in a well-understood process resulting in ringdown radiation, will allow us to test the "no-hair" theorem for black holes (Israel, 1967). Another test would be to verify the Hawking area theorem (Hawking, 1973), by determining that the area of the final black hole exceeds the sum of the areas of the initial holes, as determined by the fits to their masses and spins. The area theorem should be satisfied in most circumstances if GR is correct and the objects are black holes, but might be violated if the central objects are made of some exotic matter.

[1] The expression "standard sirens" has been chosen in analogy to the standard candles in the electromagnetic spectrum.

3.3. Characteristics of the signals

The gravitational wave signal from the MBHB can be conventionally split into three parts: inspiral, merger and the ringdown. The inspiral part of the waveform is also referred to as a *chirp*. Its amplitude and frequency increase as two body spiral toward each other due to loss in energy and angular momentum from the system caused by the emission of gravitational waves. The signal depends on masses, source's location on the sky and inclination of the orbital plane to the line of sight. The duration of the inspiral in the LISA's band decreases with increasing total mass and, for a fixed total mass, the duration is longer for unequal mass binaries: $t \sim M^{-5/3}\eta^{-1}$, where M is a total mass and $\eta = m_1 m_2/M^2$. The inspiral signal is broadband and its bandwidth is inversly proportional to the total mass of the system. For the high mass binaries we will see only the parts of the signal that correspond to the merger and the ringdown phase of the MBHB's evolution. If MBHs are spinning, then spin-spin and spin-orbital interactions will cause both amplitude and phase modulation of the signal which are increasingly important for unequal mass binaries. For unequal mass binaries, the orbital eccentricity could also be measurable as they enter LISA's band and it has to be taken into account. Orbits for comparable mass binaries (mass ratio smaller than 10) should be fairly circular by the time the enter observational frequency window. A detection of inspiral events alone can be expected to give SNR of $10^2 \ldots 10^3$ at $z = 1$.

The merger (*i. e.*, the last few cycles of the orbital evolution) of two compact objects and the associated final burst of gravitational radiation is a broadband signal of a very short duration. This part of the signal is very strong and could be detected alone (Baker et al., 2007a).

The merger is followed by a "ringdown" of the final black hole, through which it loses its initial distortions to become an stationary Kerr black hole. If the final mass of the black hole is greater than $\sim 10^6\,M_\odot$ the SNR from this phase ($10^3 \ldots 10^4$) will outshine the inspiral and merger of the system (Berti et al., 2006). The ringdown signal consists of a superposition of quasi-normal modes where each mode has a complex frequency, whose real part is the oscillation frequency and whose imaginary part is the inverse of the damping time. The damping time is uniquely determined by the mass and angular momentum of the newly born black hole.

Because of the nature and the strength of the sources, there will be almost no confusion between multiple MBHB signals in the same data stream, making them the cleanest sources that we have to deal with (Cornish and Porter, 2006c).

3.4. Goals of data analysis

The main goals of data analysis are to detect the MBHB systems (which is relatively easy task given the high SNR of the signal) and to measure parameters of the system:

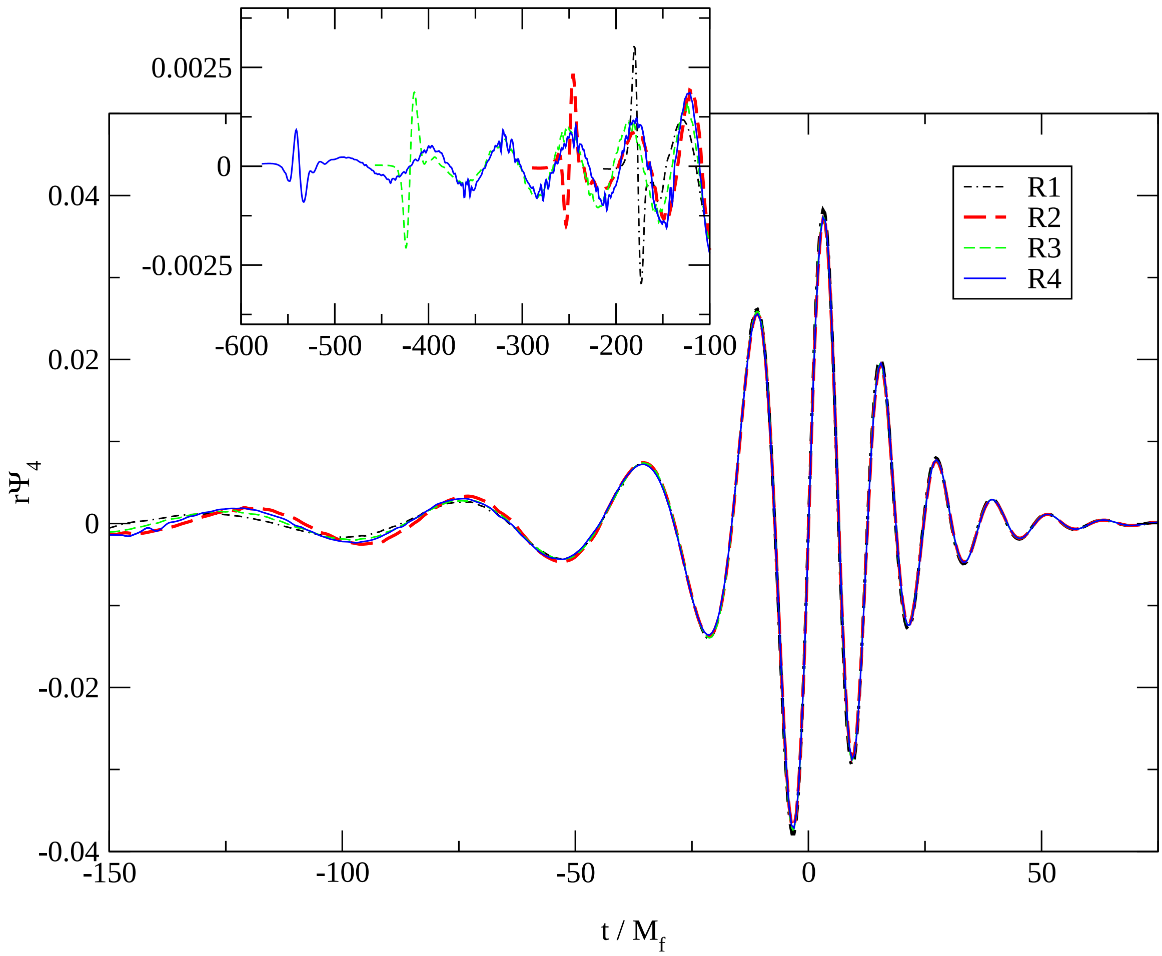


Figure 3.1.: Waveforms from numerical simulations for equal-mass MBHB starting from different initial separations, labelled R1–R4 (see Table below). The figure shows nearly perfect agreement after $t = -50\,M$, with M the final mass of the MBHB. For the preceding 500 M, shown in an inset, the waveforms agree in phase and amplitude within about 10 % except for a brief initial pulse at the beginning of each run. Graph from Baker et al. (2006).

Run	ADM mass M_0	Angular momentum J_0	Initial proper distance L/M_0
R1	0.996	0.868	9.9
R2	1.001	0.899	11.1
R3	1.002	0.928	12.1
R4	1.003	0.959	13.2

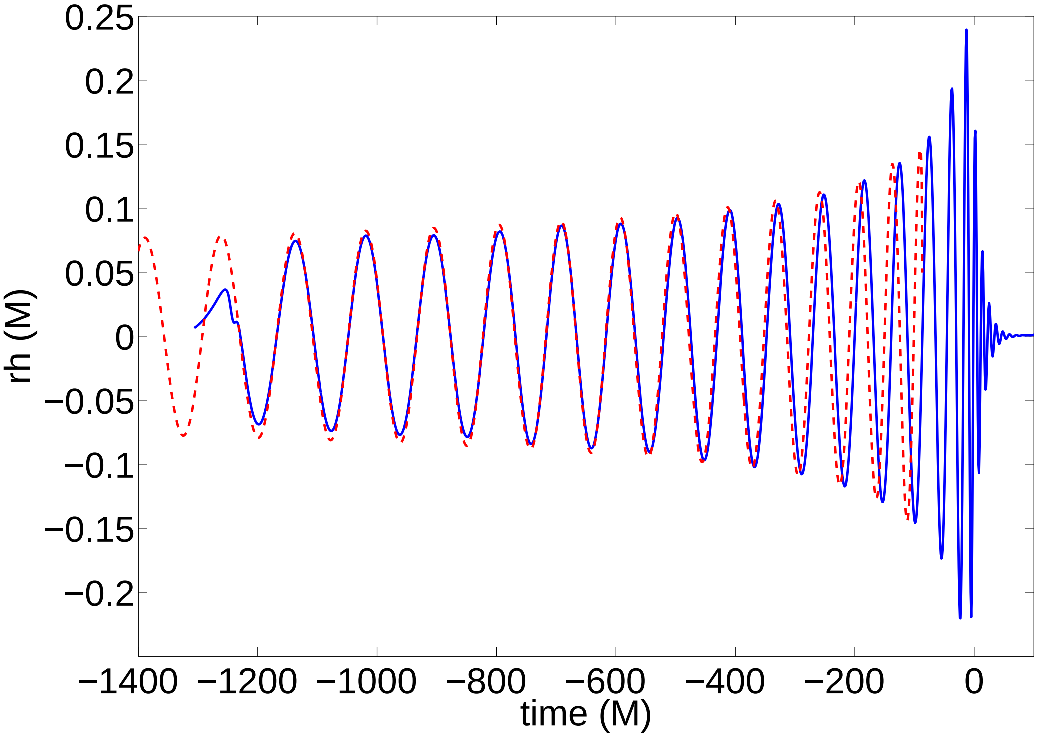


Figure 3.2.: Gravitational-strain waveforms from the merger of equal-mass Schwarzschild black holes. The solid curve is the waveform from the high-resolution numerical simulation, and the dashed curve is a PN waveform with 3.5PN order phasing and 2.5PN order amplitude accuracy. Time $t = 0$ is the moment of peak radiation amplitude in the simulation. Graph from Baker et al. (2007b)

masses, spins, and position and orientation in the sky.

One of the objectives is to detect signal using the inspiral phase alone and estimate its sky position accurately enough in order to issue warning to the astronomical community for possible simultaneous observation of merger (and post-merger) with working X-ray and optical observatories. This should enable us to identify the electromagnetic counterpart (hosting galaxy) of the gravitational wave signal.

Each mode of the ringdown signal allows an independent determination of the black holes' mass and spins (Berti et al., 2006); testing the "no-hair" theorem requires more than one quasi-normal mode frequencies in the ringdown.

3.5. Requirements of data analysis

These goals are achievable provided we have waveforms that accurately mirror nature. Such waveforms will be required to:

- Investigate the combined inspiral, merger, and ringdown signal. While it should be possible to search for each of the three phases independently, a more efficient method should be to use a combined template.
- Include the higher-PN-order and spin-induced corrections to the amplitude and inspiral phase of the waveform. While the simpler *restricted post-Newtonian* waveforms should be sufficient for detection purposes, amplitude modulations will need to be accounted for if we require accurate parameter extraction.
- Include eccentricity in the waveforms. While the system should be sufficiently circularized by time it enters the LISA bandwidth, there may be some residual eccentricity. This should be quite a small, but it may still be large enough that the assumption of circular orbits is not yet valid. A small mismatch in phase due to using the wrong waveform model could lead to a large degradation in parameter estimation.
- Be able to subtract the high SNR signals with residuals below the noise level.

3.6. Progress in development of data-analysis methods

A sufficient understanding of the expected signals can only be won from reliable GR calculations producing detailed predictions for the production of gravitational radiation. The primary tool for deriving waveform predictions for data-analysis applications, thus far, has been the post-Newtonian (PN) approximation, an expansion in powers of v/c which is very effective whenever the black holes are sufficiently well separated and the orbital velocities are comparatively small (see for instance Buonanno et al., 2003).

Waveform-phasing calculations, most crucial for matched filtering, have been carried out to order $(v/c)^7$ for non-spinning binaries in quasi-circular (Blanchet, 2006) and for eccentric orbits (Königsdörffer and Gopakumar, 2006). The spin–orbit coupling was computed for quasi-circular orbits up to 1PN order beyond the leading contribution, and the spin–spin effects are usually included at the leading (2PN) level (Blanchet et al., 2006; Faye et al., 2006). Presently, there are no post-Newtonian waveforms which include both the eccentricity of the orbit and spins of the orbiting bodies.

The higher-order correction to the amplitude of the gravitational wave were computed up to 2.5PN (Arun et al., 2006) and up to 1PN for spinning binaries (Kidder, 1995) in quasi-circular orbits. The impact of the higher orbital harmonics on data analysis was investigated by Van Den Broeck and Sengupta (2007). Those calculations cover the inspiral phase.

While we believe we have good knowledge of the inspiral and ringdown phases, the merger phase is less well understood, although there has been trmendous progress in the last few years.

An alternative approach for deriving GR-consistent waveform predictions, especially in the strong-field regime where the PN expansion fails, is numerical relativity. The basic premise of numerical relativity is to directly solve the Einstein field equations on a supercomputer.

After the first results for merging black holes were published (Baker et al., 2006; Campanelli et al., 2006; Pretorius, 2005) the field is currently advancing at a dramatic rate, allowing to accurately calculate the waveforms of non-spinning MBHB with mass ratios up to 1 : 6 (Baker et al., 2008, and references therein for earlier work of that group), of black hole binaries with fixed mass ratio but generic spin (Campanelli et al., 2008, and references therein), of binaires in unstable orbits (Pretorius and Khurana, 2007), and to assess the gravitational recoil of the system (Holley-Bockelmann et al., 2008; Lousto and Zlochower, 2008; Schnittman et al., 2008)

Comparisons of PN-waveforms with the waveforms obtained by numerical relativity have shown the consistency of the different approaches (Baker et al., 2007c; Campanelli et al., 2008; Pan et al., 2008), so there is good reason for optimism that a combination of PN and numerical relativity calculations will provide a sufficient understanding of the full coalescence waveforms for matched-filtering analysis of LISA data.

The data-analysis problem for MBHB is challenging because of the large size of the parameter space. While the data analysis for MBHB is being conducted by a small group of people at present, there have already been many breakthroughs. The Mock LISA Data Challenge (MLDC)s (*cf.* chapter 8.2) have been a major driving force in this development, with a number of groups submitting entries. While there was some overlap in the techniques used by different groups, in general each group used a variety of different algorithms.

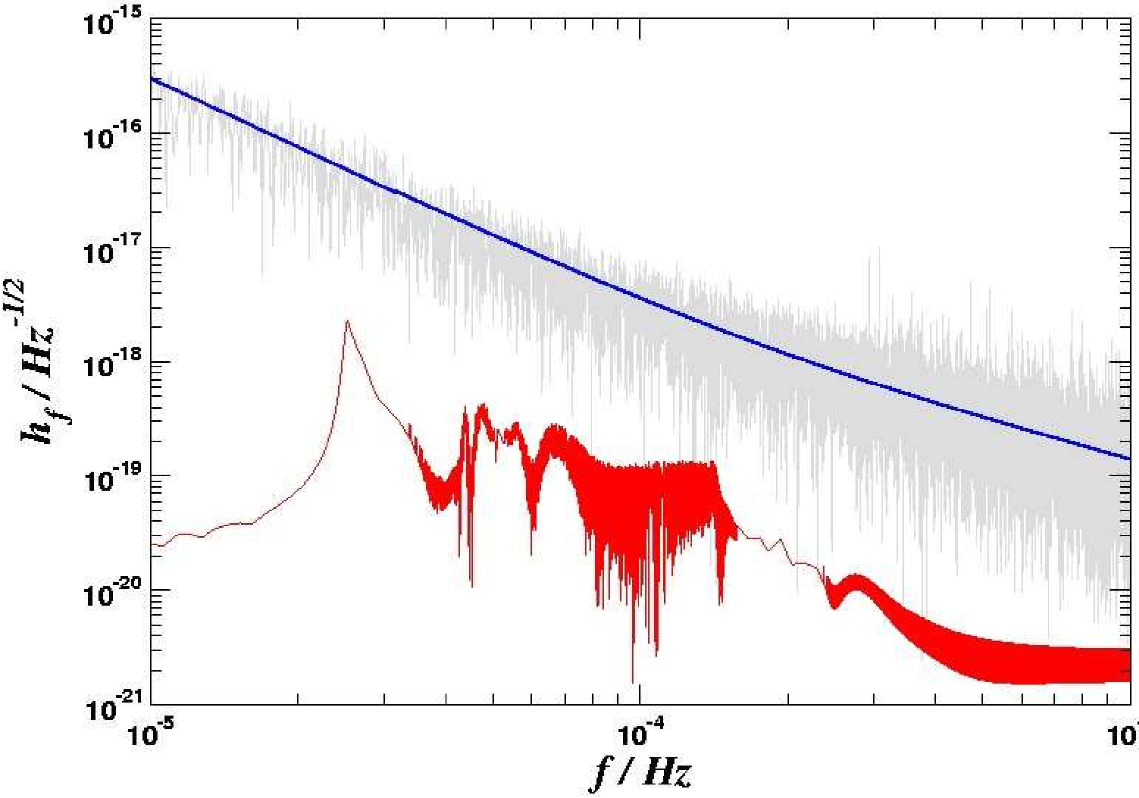


Figure 3.3.: Residuals after removing two SMBH simultaneously present in the data stream. The SMBH have redhsifts of $z = 10$ (SNR ~ 40) and $z = 1.5$ (SNR ~ 100), respectively. The data contains the simulated signals from about 27×10^6 white-dwarf–white-dwarf binaries, representing the total white-dwarf binary population in our galaxy.
The grey line shows the linear power spectral density of the combined instrumental noise and the galactic foreground, the blue line is the average per frequency bin. The red line shows the residuals after identification and removal of the two SMBH signals. Graph from Cornish and Porter (2006c).

The different algorithms include genetic algorithms, time-frequency methods, tomographic reconstruction, template bank methods and Markov Chain Monte Carlo methods (MCMC). Methods such as time-frequency and tomographic reconstruction, work as a first-stage algorithm in the context of a hierarchical search (Brown et al., 2007), whereas hierarchical template banks can be used to conduct global searches of parameter space (Harry et al., 2008). At present one of the most promising and more developed methods is the MCMC algorithm (Cornish and Porter, 2005, 2006a,b,c; Stroeer et al., 2006; Wickham et al., 2006). This is a very useful method for examining large parameter spaces, especially in the context of a large number of signals. There are many variants of the MCMC in use at the moment. However, it is clear that many groups are converging to using MCMC algorithms either as a final hierarchical stage, such as in Babak (2008); Brown et al. (2007) or as a global search algorithm.

Assuming a waveform model that is representative enough, LISA will have no problems detecting MBHB out to a redshift z of at least $z = 10$ (Cornish and Porter, 2006c). The MCMC also allows to map out the posterior distribution functions for the waveform parameters in a very easy manner.

The various groups involved in the MLDCs have shown that regardless of the method used, we should have no problems extracting MBHB sources from the LISA data stream, again assuming that our current best waveform model fully reflects nature. Because these sources are so clean it has been demonstrated that once a detection has been made, the source can be removed from the data stream with minimal contamination.

3.7. Summary and conclusions

MBHB are an attractive source for LISA data analysis. Due to their brightness and to the virtual absence of source confusion, these sources can be detected with relative ease, assuming that our current best waveform model reflects nature sufficiently well. Because these sources are so clean, once a detection has been made they can be removed from the data stream with minimal contamination. While there are still a large number of implementation aspects to be tackled, there has been good progress and as no fundamental issues prevent further development, we expect to meet all the requirements for analysis of MBHB signals.

3.8. References

K. G. Arun, B. R. Iyer, M. S. S. Qusailah, and B. S. Sathyaprakash. Letter to the editor: Testing post-Newtonian theory with gravitational wave observations. *Class. Quantum Grav.*, 23:L37–L43, May 2006, also as arXiv:gr-qc/0604018.

S. Babak. Building a stochastic template bank for detecting massive black hole binaries. *Class. Quantum Grav.*, 25(19):195011–+, October 2008, also as arXiv:0801.4070.

J. G. Baker, J. Centrella, D.-I. Choi, M. Koppitz, and J. van Meter. Binary black hole merger dynamics and waveforms. *Phys. Rev. D*, 73(10):104002–+, May 2006, also as arXiv:gr-qc/0602026.

J. G. Baker, S. T. McWilliams, J. R. van Meter, J. Centrella, D.-I. Choi, et al. Binary black hole late inspiral: Simulations for gravitational wave observations. *Phys. Rev. D*, 75(12):124024–+, June 2007a, also as arXiv:gr-qc/0612117.

J. G. Baker, J. R. van Meter, S. T. McWilliams, J. Centrella, and B. J. Kelly. Consistency of post-Newtonian waveforms with numerical relativity. *Physical Review Letters*, 99(18):181101–+, November 2007b, also as arXiv:gr-qc/0612024.

J. G. Baker, J. R. van Meter, S. T. McWilliams, J. Centrella, and B. J. Kelly. Consistency of Post-Newtonian Waveforms with Numerical Relativity. *Physical Review Letters*, 99(18):181101–+, November 2007c, also as arXiv:gr-qc/0612024.

J. G. Baker, W. D. Boggs, J. Centrella, B. J. Kelly, S. T. McWilliams, et al. Mergers of nonspinning black-hole binaries: Gravitational radiation characteristics. *Phys. Rev. D*, 78(4):044046–+, August 2008, also as arXiv:0805.1428.

E. Berti, A. Buonanno, and C. M. Will. Estimating spinning binary parameters and testing alternative theories of gravity with LISA. *Phys. Rev. D*, 71:084025, April 2005a, also as arXiv:gr-qc/0411129.

E. Berti, A. Buonanno, and C. M. Will. Testing general relativity and probing the merger history of massive black holes with LISA. *Class. Quantum Grav.*, 22: S943–S954, September 21 2005b, also as arXiv:gr-qc/0504017.

E. Berti, V. Cardoso, and C. M. Will. Gravitational-wave spectroscopy of massive black holes with the space interferometer LISA. *Phys. Rev. D*, 73:064030–+, March 2006, also as arXiv:gr-qc/0512160.

L. Blanchet, A. Buonanno, and G. Faye. Higher-order spin effects in the dynamics of compact binaries. II. Radiation field. *Phys. Rev. D*, 74(10):104034–+, November 2006, also as arXiv:gr-qc/0605140.

L. Blanchet. Gravitational Radiation from Post-Newtonian Sources and Inspiralling Compact Binaries. *Living Reviews in Relativity*, 9(4), 2006.

T. A. Boroson and T. R. Lauer. A Candidate Sub-Parsec Supermassive Binary Black Hole System. *ArXiv e-prints*, January 2009, also as arXiv:0901.3779.

D. A. Brown, J. Crowder, C. Cutler, I. Mandel, and M. Vallisneri. A three-stage search for supermassive black-hole binaries in LISA data. *Class. Quantum Grav.*, 24:595–+, October 2007, also as arXiv:0704.2447.

A. Buonanno, Y. Chen, and M. Vallisneri. Detection template families for gravitational waves from the final stages of binary black-hole inspirals: Nonspinning case. *Phys. Rev. D*, 67(2):024016–+, January 2003.

M. Campanelli, C. O. Lousto, and Y. Zlochower. Last orbit of binary black holes. *Phys. Rev. D*, 73(6):061501–+, March 2006, also as arXiv:gr-qc/0601091.

M. Campanelli, C. O. Lousto, H. Nakano, and Y. Zlochower. Comparison of Numerical and Post-Newtonian Waveforms for Generic Precessing Black-Hole Binaries. *ArXiv e-prints*, August 2008, also as arXiv:0808.0713.

N. J. Cornish and E. K. Porter. Detecting galactic binaries with LISA. *Class. Quantum Grav.*, 22:S927–S933, September 21 2005, also as arXiv:gr-qc/0504012.

N. J. Cornish and E. K. Porter. Catching Super Massive Black Hole Binaries Without a Net. *ArXiv General Relativity and Quantum Cosmology e-prints*, May 2006a, also as arXiv:gr-qc/0605135.

N. J. Cornish and E. K. Porter. MCMC exploration of supermassive black hole binary inspirals. *Class. Quantum Grav.*, 23:761–+, October 2006b.

N. J. Cornish and E. K. Porter. The Search for SuperMassive Black Hole Binaries with LISA. *ArXiv General Relativity and Quantum Cosmology e-prints*, December 2006c, also as arXiv:gr-qc/0612091.

N. Dalal, D. E. Holz, S. A. Hughes, and B. Jain. Short GRB and binary black hole standard sirens as a probe of dark energy. *Phys. Rev. D*, 74(6):063006, September 2006, also as arXiv:astro-ph/0601275.

C. Deffayet and K. Menou. Probing Gravity with Spacetime Sirens. *ApJ*, 668: L143–L146, October 2007, also as arXiv:0709.0003.

G. Faye, L. Blanchet, and A. Buonanno. Higher-order spin effects in the dynamics of compact binaries. I. Equations of motion. *Phys. Rev. D*, 74(10):104033–+, November 2006, also as arXiv:gr-qc/0605139.

M. G. Haehnelt. Hierarchical build-up of galactic bulges and the merging rate of supermassive binary black holes. *Class. Quantum Grav.*, 20:31–+, May 2003, also as arXiv:astro-ph/0307379.

I. W. Harry, S. Fairhurst, and B. S. Sathyaprakash. A hierarchical search for gravitational waves from supermassive black hole binary mergers. *Class. Quantum Grav.*, 25(18):184027–+, September 2008, also as arXiv:0804.3274.

S. W. Hawking. A variational principle for black holes. *Communications in Mathematical Physics*, 33:323–334, December 1973.

K. Holley-Bockelmann, K. Gültekin, D. Shoemaker, and N. Yunes. Gravitational Wave Recoil and the Retention of Intermediate-Mass Black Holes. *ApJ*, 686: 829–837, October 2008, also as arXiv:0707.1334.

D. E. Holz and S. A. Hughes. Using Gravitational-Wave Standard Sirens. *ApJ*, 629: 15–22, August 2005, also as arXiv:astro-ph/0504616.

P. F. Hopkins, L. Hernquist, T. J. Cox, T. Di Matteo, B. Robertson, et al. A Unified, Merger-driven Model of the Origin of Starbursts, Quasars, the Cosmic X-Ray Background, Supermassive Black Holes, and Galaxy Spheroids. *ApJS*, 163:1–49, March 2006, also as arXiv:astro-ph/0506398.

W. Israel. Event Horizons in Static Vacuum Space-Times. *Physical Review*, 164: 1776–1779, December 1967.

L. E. Kidder. Coalescing binary systems of compact objects to (post)$^{5/2}$-Newtonian order. V. Spin effects. *Phys. Rev. D*, 52:821–847, July 1995, also as arXiv:gr-qc/9506022.

B. Kocsis, Z. Frei, Z. Haiman, and K. Menou. Finding the Electromagnetic Counterparts of Cosmological Standard Sirens. *ApJ*, 637:27–37, January 2006, also as arXiv:astro-ph/0505394.

C. Königsdörffer and A. Gopakumar. Phasing of gravitational waves from inspiralling eccentric binaries at the third-and-a-half post-Newtonian order. *Phys. Rev. D*, 73(12):124012–+, June 2006, also as arXiv:gr-qc/0603056.

E. V. Linder. Gravitational wave sirens as a triple probe of dark energy. *Journal of Cosmology and Astro-Particle Physics*, 3:19–+, March 2008, also as arXiv:0711.0743.

C. O. Lousto and Y. Zlochower. Further insight into gravitational recoil. *Phys. Rev. D*, 77(4):044028–+, February 2008, also as arXiv:0708.4048.

Y. Pan, A. Buonanno, J. G. Baker, J. Centrella, B. J. Kelly, et al. Data-analysis driven comparison of analytic and numerical coalescing binary waveforms: Nonspinning case. *Phys. Rev. D*, 77(2):024014–+, January 2008, also as arXiv:0704.1964.

E. Poisson. Measuring black-hole parameters and testing general relativity using gravitational-wave data from space-based interferometers. *Phys. Rev. D*, 54: 5939–5953, November 1996, also as arXiv:gr-qc/9606024.

S. F. Portegies Zwart, H. Baumgardt, S. L. W. McMillan, J. Makino, P. Hut, et al. The Ecology of Star Clusters and Intermediate-Mass Black Holes in the Galactic Bulge. *ApJ*, 641:319–326, April 2006, also as arXiv:astro-ph/0511397.

F. Pretorius. Evolution of Binary Black-Hole Spacetimes. *Physical Review Letters*, 95 (12):121101–+, September 2005, also as arXiv:gr-qc/0507014.

F. Pretorius and D. Khurana. Black hole mergers and unstable circular orbits. *Class. Quantum Grav.*, 24:83–+, June 2007, also as arXiv:gr-qc/0702084.

C. Rodriguez, G. B. Taylor, R. T. Zavala, A. B. Peck, L. K. Pollack, et al. A Compact Supermassive Binary Black Hole System. *ApJ*, 646:49–60, July 2006, also as arXiv:astro-ph/0604042.

J. D. Schnittman, A. Buonanno, J. R. van Meter, J. G. Baker, W. D. Boggs, et al. Anatomy of the binary black hole recoil: A multipolar analysis. *Phys. Rev. D*, 77 (4):044031–+, February 2008, also as arXiv:0707.0301.

A. Stroeer, J. Gair, and A. Vecchio. Automatic Bayesian inference for LISA data analysis strategies. In S. M. Merkowitz and J. C. Livas, editors, *LASER INTERFEROMETER SPACE ANTENNA: 6th International LISA Symposium*, volume 873, pages 444–451. AIP, September 2006, also as arXiv:gr-qc/0609010.

C. Van Den Broeck and A. Sengupta. Binary black hole spectroscopy. *Class. Quantum Grav.*, 24:1089–1113, March 2007, also as arXiv:gr-qc/0610126.

E. D. L. Wickham, A. Stroeer, and A. Vecchio. A Markov chain Monte Carlo approach to the study of massive black hole binary systems with LISA. *Class. Quantum Grav.*, 23:819–+, October 2006.

C. M. Will and N. Yunes. Testing alternative theories of gravity using LISA. *Class. Quantum Grav.*, 21:4367–4381, September 21 2004, also as arXiv:gr-qc/0403100.

4. Extreme and Intermediate Mass Ratio Inspirals (EMRIs and IMRIs)

4.1. Description of the sources

Supermassive black holes in the centers of galaxies regularly capture objects from the clouds of stars surrounding them. Many of these are compact stars and black holes that have, by chance, been deflected onto orbits close enough to the black hole for their first encounter with the hole to place them into a bound orbit, through the loss of energy to gravitational radiation (Gürkan and Hopman, 2007; Hopman and Alexander, 2006; Ivanov, 2002; Perets et al., 2007; Sigurdsson and Rees, 1997). Other objects may have been captured by tidal effects: compact-object binaries that have come too close to the hole and been disrupted by tidal forces (Miller et al., 2005); compact cores of giant stars that have been tidally stripped by the central hole (Di Stefano et al., 2001); or black holes that have formed near the central hole by the evolution of giant stars such as those seen near the central hole in the Milky Way (Nayakshin and Sunyaev, 2005; Paumard et al., 2006) and in the center of M31 (Bender et al., 2005). The resulting compact objects (neutron stars, black holes, white dwarfs) would then survive as point-like objects orbiting the central hole until their losses to gravitational radiation send them across the horizon. They would radiate waves in the LISA band for a long time, many of them emitting approximately 10^5 cycles or more of radiation before falling into the hole. These objects are interesting both because of what they tell us about the astrophysics of the stellar populations near central galactic black holes and because they probe the geometry of spacetime around massive black holes with unprecedented precision. A review of the sources, the astrophysics, science and detection of EMRIs using LISA can be found in Amaro-Seoane et al. (2007).

We distinguish two principal kinds of systems: EMRIs and IMRIs. EMRIs are *E*xtreme *M*ass *R*atio *I*nspiral sources, where the ratio of the mass of the inspiraling object to that of the central black hole is 10^{-4} or smaller. These objects are white dwarfs, neutron stars, or stellar-mass black holes. Because they are more massive, black holes sink to the center of star clusters and will be over-represented among the EMRI population. Also because of their mass they can be seen further away than neutron-star or white-dwarf captures, so the typical expected system is a stellar black hole with a mass of $10\,M_\odot$ falling into a $10^6\,M_\odot$ supermassive black hole. IMRIs are *I*ntermediate *M*ass *R*atio *I*nspiral sources, where the ratio is somewhat larger, between 10^{-4} and 10^{-2}. The inspiraling object here is expected to be an intermediate mass black hole, such as may have been formed in the first generation of star formation (Population III) (Miller, 2005). The canonical expected source is an intermediate-mass black hole with a mass of $10^3\,M_\odot$ falling into a $10^6\,M_\odot$ supermassive black hole.

4.2. Science return from detecting these sources

EMRIs and IMRIs figure among the principal fundamental-physics goals of the LISA mission because their signals contain rich information about the geometry around the central black holes. When they radiate in the LISA band, these objects are so close to their black holes that surrounding stars or even accretion disks will have at most a small effect on their orbits. The phase evolution of their signals, lasting for thousands or even hundreds of thousands of cycles, reflect in detail the near-geodesic orbits they follow around the holes. From this phase information we expect to measure how closely the geometry matches the Kerr geometry predicted by general relativity, and thereby test the black-hole uniqueness theorems of Einstein's theory (Barack and Cutler, 2007). Direct evidence of the existence of a horizon in the spacetime will come from seeing the signals cut off as they cross the horizon. If they do not cut off, then that will indicate that the central object is not massive black hole; explaining what it is will require exotic new physics.

There are also astrophysics payoffs for detecting these signals. EMRIs and IMRIs sample the stellar population near the central black holes. The different ways in which EMRIs might form can be studied separately because their populations will have distinguishable properties: compact objects captured directly from plunge orbits will have high eccentricity; objects that are captured by tidal effects will have nearly circular orbits; and objects that evolved near the central hole will be in circular orbits in the hole's equatorial plane, which is presumably the plane of the accretion disk. When coupled with computer models of mass segregation, stellar interactions, and population evolution, the rates and mass spectrum of the different classes of detected objects will provide unique insight into the co-evolution of black holes and their host galaxies.

Provided LISA detects a suitable number of EMRIs (about 20 to $z = 0.5$), it is possible to establish an independent redshift-distance relationship, thus determining Hubble's constant to a precision of better than 1 % (MacLeod and Hogan, 2008).

4.3. Characteristics of the signals

The frequency of EMRI and IMRI signals is determined by the mass of the central black hole, but the amplitude is determined by the mass of the infalling object.

Roughly speaking, if one can perform full matched filtering, then the square of the SNR for signals from orbits around a central black hole with a given mass will be proportional to the total energy radiated, which is expected to be around 6% of the mass of the infalling object. The SNR of a "standard" EMRI consisting of a $10\,M_\odot$ black hole falling into a $10^6\,M_\odot$ black hole, observed during its whole transit through the LISA band, will be of order 30 if the source is at redshift $z = 1$, assuming ideal matched filtering. An IMRI event generated by an infalling $10^3\,M_\odot$ black hole will be correspondingly stronger, visible with similar SNR out to $z = 20$. These events are clearly highly visible, and should be detectable even with approximate methods that do not reach the ideal matched-filtering SNRs.

Rate estimates suggest that LISA will receive some tens to hundreds of signals per year from these distances (de Freitas Pacheco et al., 2006). Given that each signal might last several months, LISA could receive ten or more overlapping signals at any time, from different locations on the sky. More distant signals will blend together into a random incoherent background that is likely to be slightly below the LISA instrumental noise, but which could, for the higher rate estimates, itself be detectable. It does not seem plausible, on current best astrophysical knowledge, that the rate could either be so high that LISA becomes completely confused by distant signals, or that it could be so low that LISA would receive no detectable signals in a 5-year mission.

Individual signal waveforms are quite complex. If the source object falls toward a spinning Kerr black hole from a direction not in the equatorial plane, it will execute a complex three-dimensional orbit. To lowest order the orbit will be a geodesic, but account must be taken of the first-order effects of the object's mass and spin, which cause the orbit to evolve and eventually to take the object inside the horizon. The phase evolution of the radiation from such an orbit contains detailed information about the trajectory of the radiating object, and hence about the geometry of the spacetime. As with the simpler coalescences of comparable-mass objects, EMRIs and IMRIs are standard candles: the amplitude of the signal coupled with its phase evolution contains enough information to estimate its luminosity distance.

Understanding the expected waveform in full requires solving the equation of motion of a small object on a Kerr background, with corrections at least to first order in the mass and spin of the object. For EMRIs the mass ratio is small enough that first-order perturbation analysis is valid over the entire orbit. Even so, this is a challenging theoretical problem in general relativity, and it has attracted considerable attention during the last ten years (Drasco and Hughes, 2006; Glampedakis and Kennefick, 2002; Hughes, 2000, 2001; Sundararajan et al., 2007, 2008). It is now known in principle how to solve the problem, and researchers are currently developing implementations that will allow efficient computations of predicted waveforms. Already there are approximations that have been demonstrated to provide matched filters good enough for detecting the currently predicted synthetic

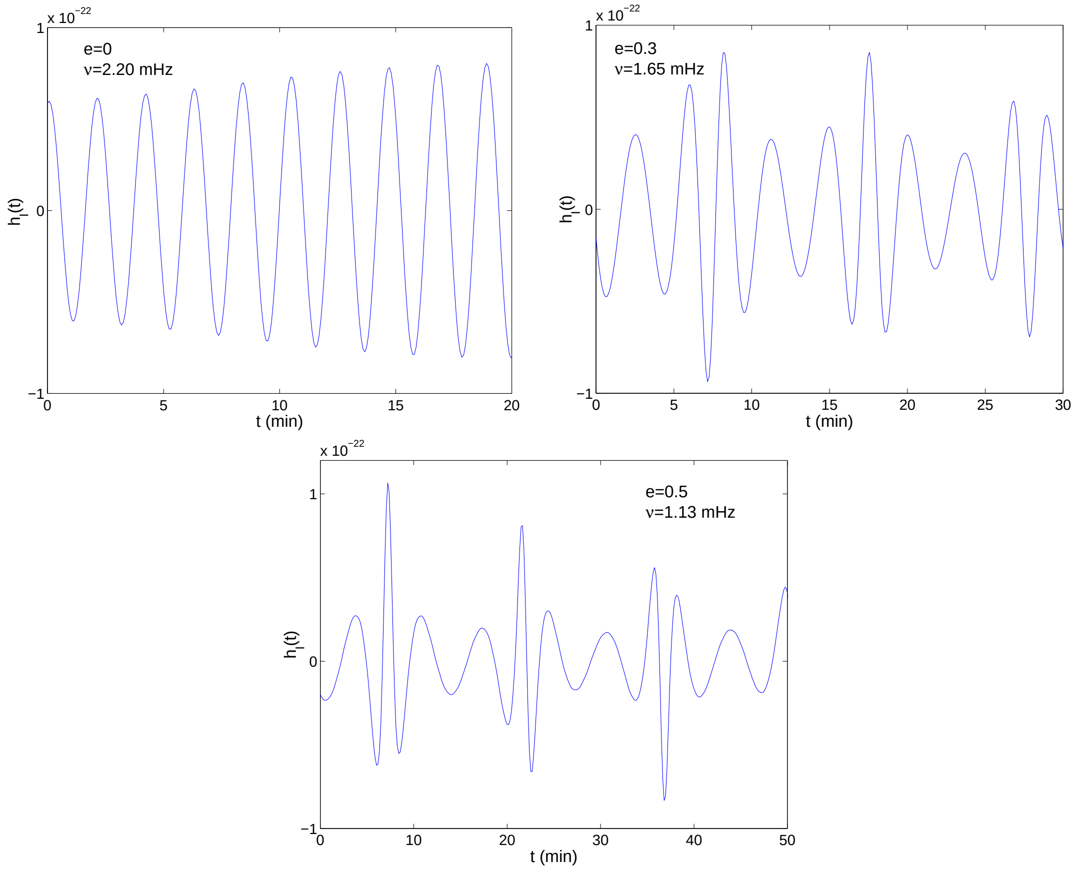


Figure 4.1.: Sample EMRI waveforms. Shown is the interferometer response function hI (t) during the last minutes before the final plunge. The three panels show cases with eccentricities at the last stable orbit of $\epsilon = 0$, $\epsilon = 0.3$, and $\epsilon = 0.5$ (top to bottom, respectively). The other physical parameters are set as follows: mass of the compact object: $\mu = 10\,M_\odot$; mass of the MBH: $M = 10^6\,M_\odot$; magnitude of the MBH's spin: $S = M^2$ (*i. e.*, maximum spin); angle between MBH's spin and orbital angular momentum: $\lambda = 30°$. Graphs from Barack and Cutler (2004)

waveforms (Babak et al., 2007; Barausse et al., 2007; Glampedakis and Babak, 2006), if not yet for extracting all the desired scientific information from them. As research in relativity develops a more complete understanding of the properties of the orbits, it is reasonable to expect that detection methods will keep pace.

For IMRIs it may be necessary to supplement perturbation theory with numerical integrations or post-Newtonian approximations in order to get a complete predicted orbit. These signals are, however, intrinsically simpler than EMRI signals, because the larger-mass objects execute fewer orbits before falling across the horizon.

A feature of EMRI and IMRI signals is the very large parameter space (17 dimensions, some of them very densely sampled) that characterizes them. The *emitted* waveform depends on the masses and spins (three-dimensional) of the central hole and the infalling object and on the orientation and eccentricity of the orbit. The infalling object's spin and the orientation and eccentricity of the orbit will change with time. In addition, the *received* signal is modulated in phase and amplitude by LISA's motion around the Sun, depending in detail on the location of the object and the orientation of LISA.

4.4. Goals of data analysis

One goal of data analysis is to be able to detect and measure strong IMRI and EMRI signals sufficiently well to test whether the central object is a black hole described by the Kerr metric, at the few percent level. This will enable fundamental tests of general relativity to be performed. A further goal is to be able to detect EMRI and IMRI signals (SNR ~ 30) and measure the masses of the inspiraling objects, the masses and spins of their central black holes, and the sky locations and distances to the systems. This will provide key information for models of the composition and evolution of the central star clusters.

4.5. Requirements for data analysis

The goals just listed for data analysis can be met if sufficiently strong signals can be detected and matched to accurately computed waveforms. The parameters of the matching signals contain the information that is required by the goals. To do this necessitates meeting the following challenging requirements:

- Construct signal templates that match the phase evolution of true signals to within one radian over periods of at least a month, and preferably for the entire duration of the signal in the LISA band. These templates will also predict amplitude and polarization evolution of the signal.

- Apply enough such templates to the LISA data stream to be able to detect reliably almost all EMRI and IMRI signals with integrated matched-filtering SNR greater than 30. The detection method must work even in the presence of overlapping signals, provided there is only one signal arriving from any single resolved direction.
- If the templates are designed to match only short (*e. g.*, one month long) stretches of the waveform, be able to infer from (from the successive templates that follow the same signal) what the important parameters of the waveform are.

4.6. Progress in development of data-analysis methods

No LISA data-analysis problem has attracted as much research effort during the last decade as the EMRI/IMRI problem. The consequence is that it is already possible to show that the first two data-analysis requirements above are solvable for the more challenging case of EMRIs, and the third one is not expected to be very difficult. Researchers have demonstrated the ability to match current synthetic waveforms for durations of order 1 month, and have shown that this is sufficient to detect longer waveforms of the required strength. Even though the parameter space of such template waveforms is very large, it has also been shown that hierarchical methods are capable of homing in on signals with SNR of 30 or more, with computing resources that are comparable to those being employed in ground-based gravitational wave projects at present, and which will be trivial to provide by the time LISA flies.

The size of the template space (the signal complexity problem) deserves some discussion. At first sight it might seem to be overwhelmingly large. A naive estimate suggests that for $\mathcal{N}$ cycles of observed gravitational wave radiation and d parameters to be fit, an optimal matched-filter search for EMRI waveforms would need $\mathcal{N}^d$ templates. As previously noted, EMRIs should emit $\sim 10^5$ cycles of radiation in the LISA band, resulting in the enormous number of $\sim 10^{5d}$ templates. Estimates of the number of important parameters in EMRI evolution range from $d \sim 7-15$; in even the most optimistic case with the smallest number of parameters, the naive expectation is that $\mathcal{N} \sim 10^{35}$ templates are needed for a search!

This would indeed be the case if the search had to find signals right down to SNR ~ 1. But we are required to look only for ideal matched-filtering SNR ~ 30 or more.[1] For strong signals, it suffices to use hierarchical methods, such as ones that use crude templates (Babak et al., 2007; Gair and Jones, 2007; Gair and Glampedakis, 2006), matching only short stretches of the signal, or that sample the template space

[1] This limit of 30 is not arbitrary. The size of the template space is so large that there is a very high chance that random noise will match some template in it. Only by setting a high threshold, certainly over 15, does one guarantee that the match is statistically significant.

very coarsely. These might find a number of false alarms as well as true signals, but the false alarms are eliminated when this first step is followed by a second filtering step that uses better templates only near the parameters of the candidate signal, to see if it really matches.

This signal complexity problem has already been encountered and addressed in ground-based searches for gravitational waves. Searches for signals from gravitational wave pulsars at kilohertz frequencies involve template spaces with some 10^{26} or more templates. The ground-based LSC community is currently doing searches that use implementations of three different semi-coherent search methods, and they will soon begin to use them in a hierarchical way. Thus, difficult as the LISA signal complexity problem is, there is reassuring legacy from ground-based analysis in addressing it.

For LISA, several detailed implementations of the first step of a hierarchical method have been investigated, such as *semi-coherent analysis* (Gair et al., 2004) and *time-frequency analysis* (Gair and Jones, 2007; Gair and Wen, 2005; Gair et al., 2008; Wen and Gair, 2005). We expect that the *Mock LISA Data Challenges* (Arnaud et al., 2006b,a; MLDC webpage) will show us what the "best" method is, but it is already clear that the problem is tractable even with currently available computing resources.

Although currently available work gives us confidence that EMRIs will be detected by LISA, there are still significant challenges to be faced if we want to reach LISA's theoretical best sensitivity (based on assuming perfect matched filtering with no source confusion). As emphasized above, the challenges involve developing ways of matching signals accurately through their whole evolution in the LISA band, and then of doing this matching against confusion from other EMRI/IMRI signals as well as against resolvable signals from binary systems in the Galaxy. There is an active research community spanning pure general relativity, signal analysis, and astrophysics that is working on this. See Table 4.1 from (Gair et al., 2004) for a guide to current expectations about rates.

There are further challenges to the community in using the detected signals and their phase evolutions to infer details of the geometry of the spacetime surrounding the central massive object, which is called the "inverse problem". Early work on the extraction of black hole maps (Ryan, 1995, 1997) showed that the multipolar information describing the central black hole's mass moments M_ℓ and mass-current moments S_ℓ is indeed encoded in the waves, and this would be enough to test whether the metric is the Kerr metric or not. The foundation for the general map extraction procedure has been laid (Brink, 2008; Glampedakis and Babak, 2006; Kesden et al., 2005). Indeed, it has been shown than an EMRI waveform allows the MBH's quadrupole moment M_2 to be determined, independently of mass and spin, to about 0.1 % (Barack and Cutler, 2007).

Table 4.1.: Number of EMRI detections. The predicted number of detectable EMRIs by LISA for a pessimistic mission scenarios (3-year mission, poor removal of confusion noise) and an optimistic mission scenarios (5-year mission, good removal of confusion noise). The rate estimates are optimistic; true rates could be a factor of up to 100 lower. The first two columns are the black hole mass $M_\bullet$ and the compact body mass m_2. Table from original data in Gair et al. (2004).

$M_\bullet$ ($10^6\,M_\odot$)	m_2 ($M_\odot$)	Optimistic #	Pessimistic #
0.3	0.6	8	0.7
0.3	10	700	89
0.3	100	1	1
1	0.6	94	9
1	10	1100	660
1	100	1	1
3	0.6	67	2
3	10	1700	134
3	100	2	1

4.7. Summary and conclusions

The EMRI/IMRI sources are among the most important sources for LISA, rich in fundamental physics and astrophysics information. They are also the most difficult data analysis problem that LISA faces, because of the large number of potentially distinguishable signals and because of the challenges in general relativity of computing reliable orbits over long durations. While these problems are not fully solved, the community has demonstrated that it is possible to extract strong signals from the data even though the signal space is large, and that it is possible to construct signal templates that are faithful to the true signals over periods as long as a month, sufficient for detection at the required level of sensitivity. Further work is needed on methods to extract full information from strong sources and the robust detection of weak sources, but good progress has been made and no fundamental obstacles encountered.

4.8. References

P. Amaro-Seoane, J. R. Gair, M. Freitag, M. C. Miller, I. Mandel, et al. TOPICAL REVIEW: Intermediate and extreme mass-ratio inspirals — astrophysics, science applications and detection using LISA. *Class. Quantum Grav.*, 24:113–+, September 2007.

K. A. Arnaud, S. Babak, J. G. Baker, M. J. Benacquista, N. J. Cornish, et al. (Mock LISA Data Challenge Task Force). A How-To for the Mock LISA Data Challenges. In S. M. Merkowitz and J. C. Livas, editors, *LASER INTERFEROMETER SPACE ANTENNA: 6th International LISA Symposium*, volume 873, pages 625–632. AIP, 2006b, also as arXiv:gr-qc/0609106.

K. A. Arnaud, S. Babak, J. G. Baker, M. J. Benacquista, N. J. Cornish, et al. (Mock LISA Data Challenge Task Force). An Overview of the Mock LISA Data Challenges. In S. M. Merkowitz and J. C. Livas, editors, *LASER INTERFEROMETER SPACE ANTENNA: 6th International LISA Symposium*, volume 873, pages 619–624. AIP, 2006a.

S. Babak, H. Fang, J. R. Gair, K. Glampedakis, and S. A. Hughes. "Kludge" gravitational waveforms for a test-body orbiting a Kerr black hole. *Phys. Rev. D*, 75(2):024005–+, January 2007, also as arXiv:gr-qc/0607007.

L. Barack and C. Cutler. LISA capture sources: Approximate waveforms, signal-to-noise ratios, and parameter estimation accuracy. *Phys. Rev. D*, 69:082005, April 15 2004, also as arXiv:gr-qc/0310125.

L. Barack and C. Cutler. Using LISA EMRI sources to test off-Kerr deviations in the geometry of massive black holes. *Phys. Rev. D*, 75(4):042003–+, February 2007, also as arXiv:gr-qc/0612029.

E. Barausse, L. Rezzolla, D. Petroff, and M. Ansorg. Gravitational waves from Extreme Mass Ratio Inspirals in non-pure Kerr spacetimes. *Phys. Rev. D*, 75(6): 064026–+, March 2007, also as arXiv:gr-qc/0612123.

R. Bender, J. Kormendy, G. Bower, R. Green, J. Thomas, et al. HST STIS Spectroscopy of the Triple Nucleus of M31: Two Nested Disks in Keplerian Rotation around a Supermassive Black Hole. *ApJ*, 631:280–300, September 2005, also as arXiv:astro-ph/0509839.

J. Brink. Spacetime encodings. I. A spacetime reconstruction problem. *Phys. Rev. D*, 78(10):102001–+, November 2008, also as arXiv:0807.1178.

J. A. de Freitas Pacheco, C. Filloux, and T. Regimbau. Capture rates of compact objects by supermassive black holes. *Phys. Rev. D*, 74(2):023001–+, July 2006, also as arXiv:astro-ph/0606427.

R. Di Stefano, J. Greiner, S. Murray, and M. Garcia. A New Way to Detect Massive Black Holes in Galaxies: The Stellar Remnants of Tidal Disruption. *ApJ*, 551: L37–L40, April 2001, also as arXiv:astro-ph/0112434.

S. Drasco and S. A. Hughes. Gravitational wave snapshots of generic extreme mass ratio inspirals. *Phys. Rev. D*, 73(2):024027, January 2006.

J. Gair and G. Jones. Detecting extreme mass ratio inspiral events in LISA data using the hierarchical algorithm for clusters and ridges (HACR). *Class. Quantum Grav.*, 24:1145–1168, March 2007, also as arXiv:gr-qc/0610046.

J. Gair and L. Wen. Detecting extreme mass ratio inspirals with LISA using time frequency methods: II. Search characterization. *Class. Quantum Grav.*, 22:S1359–S1371, September 21 2005.

J. R. Gair and K. Glampedakis. Improved approximate inspirals of test bodies into Kerr black holes. *Phys. Rev. D*, 73(6):064037–+, March 2006, also as arXiv:gr-qc/0510129.

J. R. Gair, L. Barack, T. Creighton, C. Cutler, S. L. Larson, et al. Event rate estimates for LISA extreme mass ratio capture sources. *Class. Quantum Grav.*, 21: S1595–S1606, October 21 2004.

J. R. Gair, I. Mandel, and L. Wen. Improved time frequency analysis of extreme-mass-ratio inspiral signals in mock LISA data. *Class. Quantum Grav.*, 25(18): 184031–+, September 2008, also as arXiv:0804.1084.

K. Glampedakis and S. Babak. Mapping spacetimes with LISA: inspiral of a test body in a 'quasi-Kerr' field. *Class. Quantum Grav.*, 23:4167–4188, June 21 2006, also as arXiv:gr-qc/0510057.

K. Glampedakis and D. Kennefick. Zoom and whirl: Eccentric equatorial orbits around spinning black holes and their evolution under gravitational radiation reaction. *Phys. Rev. D*, 66(4):044002–+, August 2002, also as arXiv:gr-qc/0203086.

M. A. Gürkan and C. Hopman. Resonant relaxation near a massive black hole: the dependence on eccentricity. *MNRAS*, 379:1083–1088, August 2007.

C. Hopman and T. Alexander. Resonant Relaxation near a Massive Black Hole: The Stellar Distribution and Gravitational Wave Sources. *ApJ*, 645:1152–1163, July 2006, also as arXiv:astro-ph/0601161.

S. A. Hughes. Evolution of circular, nonequatorial orbits of Kerr black holes due to gravitational-wave emission. *Phys. Rev. D*, 61(8):084004–+, April 2000.

S. A. Hughes. Evolution of circular, nonequatorial orbits of Kerr black holes due to gravitational-wave emission. II. Inspiral trajectories and gravitational waveforms. *Phys. Rev. D*, 64(6):064004–+, September 2001, also as arXiv:gr-qc/0104041.

P. B. Ivanov. On the formation rate of close binaries consisting of a super-massive black hole and a white dwarf. *MNRAS*, 336:373–381, October 2002, also as arXiv:astro-ph/0112317.

M. Kesden, J. Gair, and M. Kamionkowski. Gravitational-wave signature of an inspiral into a supermassive horizonless object. *Phys. Rev. D*, 71(4):044015, February 2005, also as arXiv:astro-ph/0411478.

C. L. MacLeod and C. J. Hogan. Precision of Hubble constant derived using black hole binary absolute distances and statistical redshift information. *Phys. Rev. D*, 77(4):043512–+, February 2008, also as arXiv:0712.0618.

M. C. Miller. Probing General Relativity with Mergers of Supermassive and Intermediate-Mass Black Holes. *ApJ*, 618:426–431, January 2005, also as arXiv:astro-ph/0409331.

M. C. Miller, M. Freitag, D. P. Hamilton, and V. M. Lauburg. Binary Encounters with Supermassive Black Holes: Zero-Eccentricity LISA Events. *ApJ*, 631:L117–L120, October 1 2005, also as arXiv:astro-ph/0507133.

MLDC webpage. Mock LISA Data Challenges, 2006. online at `http://astrogravs.nasa.gov/docs/mldc/`.

S. Nayakshin and R. Sunyaev. The 'missing' young stellar objects in the central parsec of the Galaxy: evidence for star formation in a massive accretion disc and a top-heavy initial mass function. *MNRAS*, 364:L23–L27, November 2005, also as arXiv:astro-ph/0507687.

T. Paumard, R. Genzel, F. Martins, S. Nayakshin, A. M. Beloborodov, et al. The Two Young Star Disks in the Central Parsec of the Galaxy: Properties, Dynamics, and Formation. *ApJ*, 643:1011–1035, June 2006, also as arXiv:astro-ph/0601268.

H. B. Perets, C. Hopman, and T. Alexander. Massive Perturber-driven Interactions between Stars and a Massive Black Hole. *ApJ*, 656:709–720, February 2007, also as arXiv:astro-ph/0606443.

F. D. Ryan. Gravitational waves from the inspiral of a compact object into a massive, axisymmetric body with arbitrary multipole moments. *Phys. Rev. D*, 52:5707–5718, November 1995.

F. D. Ryan. Accuracy of estimating the multipole moments of a massive body from the gravitational waves of a binary inspiral. *Phys. Rev. D*, 56:1845–1855, August 1997.

S. Sigurdsson and M. J. Rees. Capture of stellar mass compact objects by massive black holes in galactic cusps. *MNRAS*, 284:318–326, January 1997, also as arXiv:astro-ph/9608093.

P. A. Sundararajan, G. Khanna, and S. A. Hughes. Towards adiabatic waveforms for inspiral into Kerr black holes: A new model of the source for the time domain perturbation equation. *Phys. Rev. D*, 76(10):104005–+, November 2007, also as arXiv:gr-qc/0703028.

P. A. Sundararajan, G. Khanna, S. A. Hughes, and S. Drasco. Towards adiabatic waveforms for inspiral into Kerr black holes. II. Dynamical sources and generic orbits. *Phys. Rev. D*, 78(2):024022–+, July 2008, also as arXiv:0803.0317.

L. Q. Wen and J. R. Gair. Detecting extreme mass ratio inspirals with LISA using time-frequency methods. *Class. Quantum Grav.*, 22:S445–S451, May 21 2005.

5. Low-Mass Binaries (LMB)

5.1. Description of the sources

Low-mass binaries (also known as compact stellar-mass or ultra-compact binaries, LMBs for short) are binary systems containing white dwarfs, naked helium stars, neutron stars or black holes. LISA will be able to detect such binaries if they are located within our Galaxy (Nelemans et al., 2001), in the surrounding globular clusters (Benacquista, 2006) and probably even those located in some nearby galaxies (Cooray and Seto, 2005). Detection of signals from more distant extragalactic compact binaries (Farmer and Phinney, 2003) might be possible. Compact binary sources will populate the whole frequency band of the LISA detector.

The most common sources are expected to be white-dwarf–white-dwarf binaries emitting gravitational wave signals of nearly constant frequency and amplitude. Ultra-compact binaries can be classed as detached or interacting. The evolution of the detached systems is driven by gravitational radiation reaction, and the frequency evolution of these systems can be used to establish both mass and distance. In interacting systems, such as AM CVn stars, the radiation reaction competes with tides and mass transfer, and the orbital period can either increase or decrease. It has been estimated that interacting systems will contribute 13% of the total LMB signal at 1 mHz, with the contribution rising to 26% at 10 mHz (Farmer and Phinney, 2003).

From a data-analysis perspective LMBs fall into three categories: Verification Binaries, Resolved Systems, and the Confusion-limited Foreground.

5.1.1. Verification binaries

Verification binaries are the systems known from previous astronomical observations. We know sky positions and approximate periods, distances, and masses for many of these sources (see Table 5.1). Observation of these known binaries will provide a check of the operation of the instrument (Stroeer and Vecchio, 2006), as well as a strong test of GR (for example for some sources we should be able to detect or limit a scalar-field component below the 1% level). For these targeted sources matched filtering will certainly be the standard method for extraction of signals from noise.

Table 5.1.: LISA verification binaries with expected SNR larger than 5 (see Nelemans et al. (2004) and reference therein for more details). The SNR is for one interferometer and one year of integration (compiled from Stroeer and Vecchio (2006) and Roelofs et al. (2007)).

Name	f_{GW} mHz	$\dot{P}$ s/s	D pc	m_1 $M_\odot$	m_2 $M_\odot$	SNR
			AM CVn binary systems			
RXJ0806.3+1527	6.22	3.7×10^{-11}	300–1000 (?)	0.13	[0.6]	32–340
V407 Vul	3.15	3.17×10^{-12}	300–1000 (?)	0.068	[0.6]	17–126
ES Cet	3.22		350–1000 (?)	0.062	[0.6]	14–110
AM CVn	1.94		606	0.125	0.68	8–21
HP Lib	1.81		197	0.044	0.72	15–20
CR Boo	1.36		337	0.044	0.98	4
V803 Cen	1.24		347	0.057	[0.6]	4
			Ultra-compact X-ray binaries			
4U 1820-30	2.29		7600	1.4 [?]	3–10	
			Cataclysmic variables			
SDSS J1507+5230	0.49					1–5
GW Lib	0.43					1–5
WZ Sge	0.41		43	< 0.11	> 0.7	2–8

5.1.2. Resolvable binaries

The majority of the resolved systems will be found above 3 mHz. In part this is because LISA's resolving power increases with frequency, and in part because the number density of sources decreases with frequency (Bender and Hils, 1997; Hils and Bender, 2000; Hils et al., 1990). Below 3 mHz, occasional nearby and/or massive systems will stand out from crowd, permitting them to be individually resolved. It has been estimated that LISA will resolve 10^3 to 10^4 individual LMBs with one year of observations (Bender and Hils, 1997; Nelemans et al., 2001; Timpano et al., 2006).

5.1.3. Confusion-limited Foreground

At some frequency f_c, the density of sources in frequency space, (*i. e.*, the enumber of sources per frequency bin) dN/df becomes so large as to preclude the possibility of resolving all the LMBs. Exactly where this transition occurs clearly depends on the source density dN/df as well as on the efficiency of the data analysis procedure. Various arguments have been made as to the theoretically best possible performance of any data-analysis procedure. The simplest of those arguments compares the number of data points to the model dimension, which saturates at a source density of one LMB every two $1/T_{obs}$ frequency bins. Comparing this limit to the number density inferred from the various population synthesis models yields a *confusion frequency* of $f_c \simeq 3$ mHz. Below this frequency there will be an unresolved component variously described as a Confusion-limited Foreground or confusion noise.

5.2. Science return from detecting these sources

It is expected that LISA will detect and characterize around 10 000 individual compact binary systems, as well as establish an accurate estimate of the stochastic foreground produced by tens of millions of binaries in our Galaxy. The resolved systems will provide a unique map of ultra-compact binaries (for a subset of these systems both sky location and distances can be inferred). Characteristics of the population, such as the number density of sources as a function of frequency, can be used to constrain the evolutionary pathways of compact binaries.

As tracers of galactic structure, binary populations will provide observational access to many different components of the galaxy, including the disk and bulge (Seto and Cooray, 2004), the dark galactic halo (Hiscock, 1998; Hiscock et al., 2000; Ioka et al., 2000), and even globular cluster systems (Benacquista, 1999). Observations of the galactic-binary foreground with the help of gravitational waves provide a key way to measure the properties of a portion of the galactic stellar population whose presence is routinely assumed, but which has never been observed directly. The total population of highly evolved compact binaries provides a unique record of galactic stellar evolution which should be easily accessible to an instrument like LISA. Near the frequencies where LISA is most sensitive, it is also plausible that galactic-binary signals from the Milky Way's nearest companions might also be detected, and that there will be a low-level confusion background of all the galactic binaries in the local Universe (Farmer and Phinney, 2003), which would form an ultimate noise floor in the low-frequency gravitational-wave band.

5.3. Characteristics of the signals

The orbits of these binaries will have been circularized long before the observation period as a consequence of tides, common-envelope evolution, mass transfer, and emission of gravitational radiation. The signal depends on a number of parameters, some of which are intrinsic to the source, like the frequency f at which the signal is emitted, and rate of change of this frequency $\dot{f}$. Note that $\dot{f}$ can be both positive and negative. An increase in frequency, to be expected for high-frequency sources, is a consequence of the inevitable inspiral of binaries due to loss of orbital energy and angular momentum to gravitational radiation. A decrease in frequency can occur for interacting binaries with significant mass transfer or tidal coupling. In some systems the second derivative of the frequency will be measurable, and this extra information can be used to disentangle the various mechanisms that drive the frequency evolution. In some classes of LMBs, an additional parameter, namely orbital eccentricity, may be significant (Nelemans et al., 2001).

The signals from LMBs are the simplest and best understood of all the LISA sources. LMBs orbit at velocities much less than the speed of light, and the emitted signals are expected to be very accurately modelled by low-order Post-Newtonian waveforms. The motion of the LISA constellation imparts sidebands to what are otherwise essentially monochromatic signals, and the structure of the sidebands encode information about the LMBs' sky location, inclination, and orbital orientation. Other parameters that can be measured include the amplitude and the orbital phase at some fiducial time. For detached systems with large masses and/or frequencies it will be possible to infer the distance by measuring the first time derivative of the frequency.

5.4. Goals of data analysis

The primary objective for LMB data analysis is to reliably resolve and characterize several thousand individual binary systems down to SNR = 10, thereby providing a rich LMB census and reducing the level of confusion noise. The parameters to be determined include the frequency, sky location, amplitude, orbital orientation, and orbital phase. For a subset of systems the first and second frequency derivatives and the orbital eccentricity may be measurable. A further goal is to determine the power spectrum of the unresolved component and determine if the anisotropy of the foreground is consistent with a galactic origin.

5.5. Requirements for data analysis

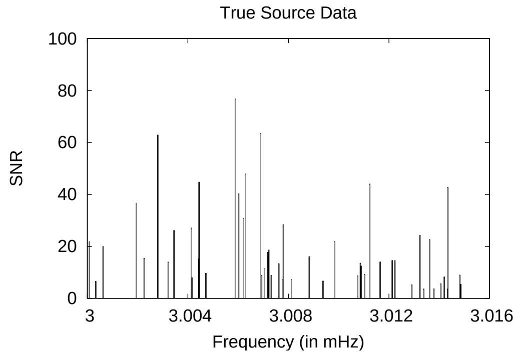
The signals from verification binaries shall be detected and parameters recovered to an accuracy that allows for a significant null test of General Relativity. The frequency, sky position, amplitude, orbital orientation and orbital phase shall be determined for a few thousand LMBs and the first and second time derivatives of the frequency evolution should be determined for hundreds and tens of LMBs, respectively. Eccentricities greater than 0.1 shall be detected for LMBs with an SNR larger than 50. The power spectrum of the unresolved galactic foreground shall be measured, and estimates given for the first few even moments of the spatial distribution.

5.6. Progress in development of data analysis methods

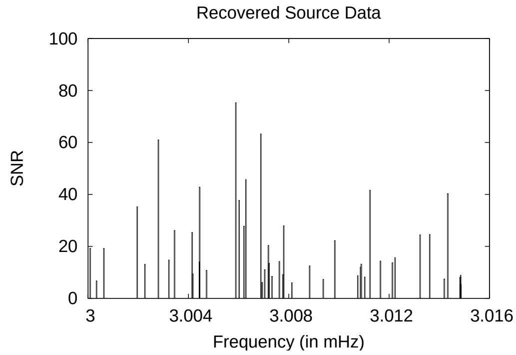
There has been considerable progress in the data analysis of LMBs over the past few years. Most approaches employ some variant of matched filtering as the templates for LMB signals are well understood and computationally inexpensive to generate, however, other approaches, such as tomographic reconstruction (Hayama et al., 2006; Mohanty and Nayak, 2006; Nayak et al., 2006) have been proposed.

In terms of detecting the signal from an isolated LMB, there is considerable overlap with the techniques employed in ground-based GW data analysis, notably in the search for (almost) monochromatic signals of the kind that would be produced by a rapidly rotating asymmetric neutron star (Brady and Creighton, 2000; Brady et al., 1998). The detection statistic employed in such a search is called the $\mathcal{F}$-statistic (Jaranowski et al., 1998); essentially, it involves maximizing the likelihood function analytically over inclination, distance, initial phase, and polarization angle. Thus we need to search only over the GW frequency, the rate of frequency change, and the position in the sky of the source.

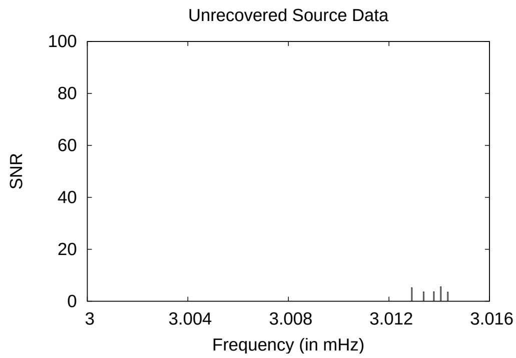
The most straightforward search method (at least conceptually) is based on constructing a grid of uniformly spaced templates (Owen, 1996). This involves computing the metric on the parameter space, choosing an optimal coordinate frame, and developing an algorithm to place the grid points so that they are separated by a constant distance called the *Minimal match*. The construction of the Minimal match is such that the inevitable reduction in SNR (and corresponding detection loss) due to coarseness of the grid remains tolerably small. It was shown (Cornish and Porter, 2005) that a search for a single LMB requires less than 10^8 templates



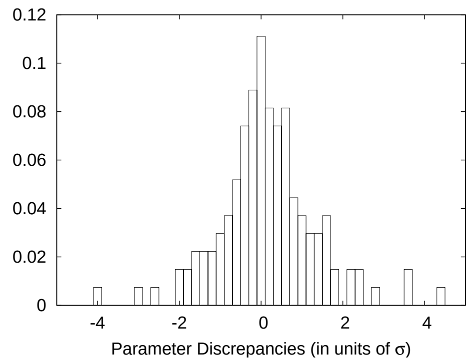
(a) Source data



(b) Recovered sources



(c) Unrecovered Sources



(d) Histogram of the parameter discrepancies

Figure 5.1.: Performance of the blocked-anneald Metropolis Hastings algorithm (BAM) on training data for the first round of the MLDC. Figure 5.1a shows the SNR and the frequency of the original sources in the training data, figure 5.1b the SNR and frequency of the recovered sources. The missed sources (*i. e.* false negatives) are shown in figure 5.1c, having an SNR of about 5 or less. Figure 5.1d shows the histogram of the discrepancies between the true intrinsic source parameters and the intrinsic parameters recovered by the BAM algorithm searching the training data set. Differences are given in units of the parameter variances. Graphs from Crowder and Cornish (2006).

below 1.6 mHz, and a few times 10^9 templates for higher frequencies. The data is filtered through each template in the grid and all instances of threshold crossings of the $\mathcal{F}$-statistic—corresponding to a match—are registered. As a search of the complete (albeit discretized) parameter space, it follows that in terms of computation effort, this method is quite expensive. To reduce computational cost this method can be applied hierarchically, starting with a coarse grid and then zooming in on the candidates.

The $\mathcal{F}$-statistic for the LISA "network" of detectors was derived by Królak et al. (2004). They also demonstrated the application of the grid-based method to simulated LISA data with one or two signals present in the data stream. Extending a grid-based search to multiple overlapping LMBs comes with a significant computational cost as the number of templates in the grid grows geometrically with the number of parameters in the model. Two approaches have been suggested for keeping the computational costs in check: iterative grid-based searches and global gridless searches (*cf.* section 2.3).

Iterative approaches such as gCLEAN (Cornish and Larson, 2003) and "Slice & Dice" (Rubbo et al., 2006) combine a grid based search for a single LMB signal with a procedure for taking into account the overlap with other sources. These methods have been applied to data streams containing a few overlapping sources. Similarily, code developed for the search for signals from rotating neutron stars in ground-based detectors was used (Whelan et al., 2008)

A very different approach is to abandon the grid in favor of what might be described as semi-random walks through parameter space. Examples include Markov Chain Monte Carlo (MCMC) Algorithms and Genetic Algorithms. The advantage of these methods is that they avoid the highly non-trivial task of laying out a uniform search grid, and the search cost grows linearly with the model dimension. The disadvantage is that there is no guarantee that the search will find the optimal fit to the data. The MCMC approach has been used in data analysis of ground based GW detector data for inspiralling neutron star binaries (Christensen and Meyer, 1998), for rapidly spinning single neutron stars (Christensen et al., 2004) and in the framework of the MLDC for LMB(Trias et al., 2008).

The Reverse Jump MCMC algorithm has been applied to a toy LISA LMB problem of extracting multiple sinusoids from Gaussian noise (Andrieu and Doucet, 1999; Umstätter et al., 2005), where it was shown that the method can detect many signals simultaneously, estimate their parameters and the instrument noise level and, very importantly, estimate the number of signals present in the data.

A variant of the MCMC approach was first applied to simulated LISA data streams with tens of overlapping sources (Cornish and Crowder, 2005), and the algorithm has since been generalized to handle the entire LISA band (Crowder and Cornish, 2006).

Different methods for computing Bayes factors have been explored, including reverse jump Markov chain Monte Carlo algorithm, Savage-Dickie density ratios, the Schwarz-Bayes information criterion, and the Laplace approximation to the model evidence (Cornish and Littenberg, 2007) and good agreement between those approaches has been found.

Genetic Algorithms are another method to handle efficiently the search for multiple sources. The method was successfully applied to parameter extraction of multiple overlapping sources in simulated LISA data streams (Crowder et al., 2006) as well as in searching for extra-galactic binaries (Crowder and Cornish, 2007). Because of their efficient searching capabilities it was suggested to use genetic algorithms as a first step in a hierarchical procedure followed by MCMC-based search.

Hybrid approaches, that combine grid-based searches for pre-selection with a MCMC algorithm for final parameter extraction have been successfully employed (Stroeer et al., 2007)

The galactic foreground can also be treated as an anisotropic, stochastic signal, by considering either the full galactic foreground, or just the residualst after the resolved systems have been taken out of the the data. In either case, the challenge is to distinguish between instrument noise and galactic confusion noise. One approach is to use the symmetric Sagnac channel (*cf.* app. A.1) to estimate the instrument noise level, then attribute any excess to galactic and extra-galactic sources (Mohanty and Nayak, 2006; Nayak et al., 2006). Another approach is to exploit the anisotropy of the foreground signal and the resulting time dependent modulation of the total galactic signal (Edlund et al., 2005a,b). See section 6 for a description of both of these approaches.

5.7. Summary and conclusions

Techniques for handling the vast number of LMB signals that are expected to populate the LISA band are now fairly mature. Techniques have been developed to simultaneously detect and characterize tens of thousands of LMB systems down to a signal to noise threshold of SNR ~ 5 and source densities as high as one source every four frequency bins (Crowder and Cornish, 2006). Methods have also been developed to isolate and characterize the unresolved galactic foreground. End-to-end demonstrations of these capabilities are one of the main goals of the second round of Mock LISA Data Challenge (MLDC)s (*cf.* chapter 8.2).

5.8. References

C. Andrieu and A. Doucet. Joint Bayesian model selection and estimation of noisy sinusoids via reversible jump MCMC. *IEEE Trans. Signal Process.*, 47(10):

2667–2676, October 1999.

M. Benacquista. Gravitational Radiation from Globular Clusters. *ApJ*, 520:233–238, July 1999, also as arXiv:astro-ph/9810484.

M. J. Benacquista. Relativistic Binaries in Globular Clusters. *Living Reviews in Relativity*, 9(2), February 2006.

P. L. Bender and D. Hils. Confusion noise level due to galactic and extragalactic binaries. *Class. Quantum Grav.*, 14:1439–1444, June 1997.

P. R. Brady and T. Creighton. Searching for periodic sources with LIGO. II. Hierarchical searches. *Phys. Rev. D*, 61(8):082001–+, April 2000, also as arXiv:gr-qc/9812014.

P. R. Brady, T. Creighton, C. Cutler, and B. F. Schutz. Searching for periodic sources with LIGO. *Phys. Rev. D*, 57:2101–2116, February 1998, also as arXiv:gr-qc/9702050.

N. Christensen and R. Meyer. Markov chain Monte Carlo methods for Bayesian gravitational radiation data analysis. *Phys. Rev. D*, 58(8):082001–+, October 1998.

N. Christensen, R. J. Dupuis, G. Woan, and R. Meyer. Metropolis-Hastings algorithm for extracting periodic gravitational wave signals from laser interferometric detector data. *Phys. Rev. D*, 70(2):022001–+, July 2004, also as arXiv:gr-qc/0402038.

A. Cooray and N. Seto. Can the Laser Interferometer Space Antenna Resolve the Distance to the Large Magellanic Cloud? *ApJ*, 623:L113–L116, April 20 2005, also as arXiv:astro-ph/0502500.

N. J. Cornish and J. Crowder. LISA data analysis using Markov chain Monte Carlo methods. *Phys. Rev. D*, 72:043005, August 2005, also as arXiv:gr-qc/0506059.

N. J. Cornish and S. L. Larson. LISA data analysis: Source identification and subtraction. *Phys. Rev. D*, 67:103001–+, May 15 2003, also as arXiv:astro-ph/0301548.

N. J. Cornish and T. B. Littenberg. Tests of Bayesian model selection techniques for gravitational wave astronomy. *Phys. Rev. D*, 76(8):083006–+, October 2007, also as arXiv:0704.1808.

N. J. Cornish and E. K. Porter. Detecting galactic binaries with LISA. *Class. Quantum Grav.*, 22:S927–S933, September 21 2005, also as arXiv:gr-qc/0504012.

J. Crowder and N. Cornish. A Solution to the Galactic Foreground Problem for LISA. *ArXiv Astrophysics e-prints*, November 2006, also as arXiv:astro-ph/0611546.

J. Crowder and N. J. Cornish. Extracting galactic binary signals from the first round of Mock LISA Data Challenges. *Class. Quantum Grav.*, 24:575–+, October 2007, also as arXiv:0704.2917.

J. Crowder, N. J. Cornish, and J. L. Reddinger. LISA data analysis using genetic algorithms. *Phys. Rev. D*, 73:063011–+, March 2006, also as arXiv:gr-qc/0601036.

J. A. Edlund, M. Tinto, A. Królak, and G. Nelemans. Simulation of the white dwarf white dwarf galactic background in the LISA data. *Class. Quantum Grav.*, 22: S913–S926, September 21 2005a, also as arXiv:gr-qc/0504026.

J. A. Edlund, M. Tinto, A. Królak, and G. Nelemans. White-dwarf white-dwarf galactic background in the LISA data. *Phys. Rev. D*, 71:024015–+, June 2005b, also as arXiv:gr-qc/0504112.

A. J. Farmer and E. S. Phinney. The gravitational wave background from cosmological compact binaries. *MNRAS*, 346:1197–1214, December 2003, also as arXiv:astro-ph/0304393.

K. Hayama, S. D. Mohanty, and K. R. Nayak. A modified CLEAN method and its application to tomographic reconstruction of LISA Galactic binaries. In S. M. Merkowitz and J. C. Livas, editors, *LASER INTERFEROMETER SPACE ANTENNA: 6th International LISA Symposium*, volume 873, pages 465–470. AIP, 2006.

D. Hils and P. L. Bender. Gravitational Radiation from Helium Cataclysmics. *ApJ*, 537:334–341, July 2000.

D. Hils, P. L. Bender, and R. F. Webbink. Gravitational radiation from the Galaxy. *ApJ*, 360:75–94, September 1990.

W. A. Hiscock. Low-Frequency Gravitational Waves from Black Hole MACHO Binaries. *ApJ*, 509:L101–L104, December 1998, also as arXiv:gr-qc/9809018.

W. A. Hiscock, S. L. Larson, J. R. Routzahn, and B. Kulick. Low-Frequency Gravitational Waves from White Dwarf MACHO Binaries. *ApJ*, 540:L5–L8, September 2000, also as arXiv:gr-qc/0005134.

K. Ioka, T. Tanaka, and T. Nakamura. Can the Density Distribution and Shape of the Galactic Dark Halo Be Determined by Low-Frequency Gravitational Waves? *ApJ*, 528:51–55, January 2000, also as arXiv:astro-ph/9903011.

P. Jaranowski, A. Królak, and B. F. Schutz. Data analysis of gravitational-wave signals from spinning neutron stars: The signal and its detection. *Phys. Rev. D*, 58 (6):063001–+, September 1998, also as arXiv:gr-qc/9804014.

A. Królak, M. Tinto, and M. Vallisneri. Optimal filtering of the LISA data. *Phys. Rev. D*, 70:022003, July 2004.

S. D. Mohanty and R. K. Nayak. Tomographic approach to resolving the distribution of LISA Galactic binaries. *Phys. Rev. D*, 73:083006–+, April 2006, also as arXiv:gr-qc/0512014.

K. R. Nayak, S. D. Mohanty, and K. Hayama. Tomographic method for resolving the Galactic binaries: including multiple interferometers and antenna patterns. In S. M. Merkowitz and J. C. Livas, editors, *LASER INTERFEROMETER SPACE ANTENNA: 6th International LISA Symposium*, volume 873, pages 429–436. AIP, 2006.

G. Nelemans, L. R. Yungelson, and S. F. Portegies Zwart. The gravitational wave signal from the Galactic disk population of binaries containing two compact objects. *A&A*, 375:890–898, September 2001, also as arXiv:astro-ph/0105221.

G. Nelemans, L. R. Yungelson, and S. F. Portegies Zwart. Short-period AM CVn systems as optical, X-ray and gravitational-wave sources. *MNRAS*, 349:181–192, March 2004, also as arXiv:astro-ph/0312193.

B. J. Owen. Search templates for gravitational waves from inspiraling binaries: Choice of template spacing. *Phys. Rev. D*, 53:6749–6761, June 1996, also as arXiv:gr-qc/9511032.

G. H. A. Roelofs, P. J. Groot, G. F. Benedict, B. E. McArthur, D. Steeghs, et al. Hubble Space Telescope Parallaxes of AM CVn Stars and Astrophysical Consequences. *The Astrophysical Journal*, 666(2):1174–1188, 2007.

L. J. Rubbo, N. J. Cornish, and R. W. Hellings. Slice & Dice: Identifying and Removing Bright Galactic Binaries from LISA Data. In S. M. Merkowitz and J. C. Livas, editors, *LASER INTERFEROMETER SPACE ANTENNA: 6th International LISA Symposium*, volume 873, pages 489–493. AIP, August 2006, also as arXiv:gr-qc/0608112.

N. Seto and A. Cooray. LISA measurement of gravitational wave background anisotropy: Hexadecapole moment via a correlation analysis. *Phys. Rev. D*, 70: 123005–+, December 2004, also as arXiv:astro-ph/0403259.

A. Stroeer and A. Vecchio. The LISA verification binaries. *Class. Quantum Grav.*, 23: 809–+, October 2006.

A. Stroeer, J. Veitch, C. Röver, E. Bloomer, J. Clark, et al. Inference on white dwarf binary systems using the first round Mock LISA Data Challenges data sets. *Classical and Quantum Gravity*, 24:541–+, October 2007, also as arXiv:0704.0048.

S. E. Timpano, L. J. Rubbo, and N. J. Cornish. Characterizing the galactic gravitational wave background with LISA. *Phys. Rev. D*, 73:122001, June 2006.

M. Trias, A. Vecchio, and J. Veitch. Markov chain Monte Carlo searches for galactic binaries in Mock LISA Data Challenge 1B data sets. *Class. Quantum Grav.*, 25(18): 184028–+, September 2008, also as arXiv:0804.4029.

R. Umstätter, N. Christensen, M. Hendry, R. Meyer, V. Simha, et al. Bayesian modeling of source confusion in LISA data. *Phys. Rev. D*, 72:022001, July 2005, also as arXiv:gr-qc/0503121.

J. T. Whelan, R. Prix, and D. Khurana. Improved search for galactic white-dwarf binaries in Mock LISA Data Challenge 1B using an {\cal F} -statistic template bank. *Class. Quantum Grav.*, 25(18):184029–+, September 2008, also as arXiv:0805.1972.

6. Stochastic signals

Stochastic signals are superpositions of large numbers of unresolved signals, typically generated by particular populations of sources. In electromagnetic observations these are often called "diffuse backgrounds". There is a strong theoretical expectation that the early universe produced stochastic gravitational waves in a broad spectrum that ranges from the longest wavelengths (currently being searched for in the "B-modes" of the cosmic microwave background temperature fluctuations), through the LISA band, and up to and beyond the ground-based band (Maggiore, 2000).

In the LISA frequency band there are in fact a large number of expected stochastic signals. They include a Galactic foreground produced by the incoherent superposition of radiation generated mainly by white-dwarf binary systems in the Galaxy; extragalactic foregrounds generated by populations of binaries in other galaxies (up to medium-to-high redshift) (Barack and Cutler, 2004; Bender and Hils, 1997; Farmer and Phinney, 2003; Hils and Bender, 2000; Hils et al., 1990) and compact objects in accretion disks of AGN (Sigl et al., 2007) ; and cosmological stochastic backgrounds produced in the early Universe (Hogan, 2006b; Maggiore, 2000; Price and Siemens, 2008). Although these signals were initially considered more as a source of "noise" in observations of individual sources rather than a new signal class, their importance for the LISA mission is now fully recognized.

The Galactic foreground is a guaranteed signal for LISA. Cosmological backgrounds, despite being the least certain of all source for the mission, might provide one of LISA's major discoveries.

The strength of a gravitational-wave background is typically characterized by the amount of energy it contains, rather than by a typical stochastic amplitude. The energy spectrum (energy per unit frequency) $\rho_{gw}(f)$ is conventionally normalized to the closure energy density of the universe, ρ_c, by defining the fraction of closure density:

$$\Omega_{gw}(f) = \frac{f}{\rho_c}\frac{d\rho_{gw}(f)}{df}.$$

Ground-based gravitational-wave projects have put considerable effort into searching for stochastic backgrounds (Abbott et al., 2005; Abbott et al., 2007). The method they employ is a cross-correlation between two independent detectors, looking for a common "noise" (the gravitational-wave field) underneath the presumably uncorrelated instrumental noise of the two instruments (Allen and Romano, 1999). This method is not available to LISA, as we shall see below. However, recent efforts dedicated to producing gravitational-wave maps (Allen and Ottewill, 1997; Ballmer, 2006) is directly applicable to the search for anisotropy in LISA data, so the amount of legacy knowledge that LISA can import from ground-based activities is growing.

6.1. Description of the sources

The stochastic signal from the Galaxy is a confusion signal consisting of the incoherent superposition of large numbers of compact-object binary systems (largely white-dwarf binaries). The definition of "incoherent" is that the sources are not resolved by LISA, which means, roughly speaking, that in a mission of duration T, the number of binary systems in a band $\Delta f = 1/T$ around any frequency f is equal to or greater than about 1. Population studies of the binaries in the Galaxy suggest that below about 1 mHz the binary signal will be stochastic (Edlund et al., 2005a; Nelemans et al., 2001). Above this limit most binaries will be resolved. Below this there may be nearby strong signals that can be resolved because they stand above the background. The "edge" of this background is not sharp, and it will be an important objective of the data analysis to resolve as many sources as possible in this edge: that is, to dig as deeply as possible into the confusion background. Observations of the resolved population will allow us to characterize the low-frequency confusion background better. Observations of the background itself will contain a wealth of information about the compact-binary population and its evolution.

Besides its sharp falloff with frequency, the Galactic background is distinguished from other backgrounds in that it is *anisotropic*. It will be stronger in the galactic plane and especially so in the direction of the galactic center rather than in orthogonal directions.

The primary stochastic signal from external galaxies comes again from compact-object binaries (Farmer and Phinney, 2003), but it is lower than the Galactic component below 1 mHz. Its exact strength depends on the population statistics in other galaxies. It may well lie below LISA's instrumental noise, but it may also be larger; most importantly, it will be isotropic.

Another stochastic signal may come from a confusion background of EMRI signals. EMRI signals from further than $z = 1$ will be too weak for LISA to detect individually, but there will be many such events superposed in LISA's data stream (Barack and Cutler, 2004). The strength of the background depends on the redshift evolution of the EMRI population. Estimates of this background suggest that it is likely to be a little below LISA's noise, but it is again possible that it is somewhat stronger. As with the Galactic binaries, this background has no sharp edge, and it will be an important challenge to the data analysis to dig as deeply as possible into it for resolved sources.

Physically the most interesting stochastic background is a cosmological background from the very early universe (Hogan, 2006b; Maggiore, 2000). Inflation predicts that tensor perturbations should be amplified by the exponential expansion of the universe in the same way as scalar perturbations (Guth, 1981; Linde, 1982; Starobinskiǐ, 1979). The scalar perturbations led to the temperature fluctuations measured by COBE, WMAP and other cosmic microwave background (CMB) experiments (Spergel et al., 2007). The ESA Planck mission will conduct a sensitive search for the tensor perturbations through the effects they have on the polarization of the cosmic microwave background (The Planck Consortium, 2005). This will reveal, or at least set limits on, the gravitational-wave background at very low frequencies, near $f = H_0$ (Chongchitnan and Efstathiou, 2006). In the LISA frequency band, standard inflation scenarios predict a very low level of background radiation, not larger than $\Omega_{gw} \sim 10^{-15}$ (Chongchitnan and Efstathiou, 2006). This is much lower than LISA's expected sensitivity of $\Omega_{gw} \sim 10^{-10}$. If LISA sees such a background, then it will clearly have important implications for the inflation scenario.

Another possible source of cosmological gravitational waves is from density perturbations that arise from inhomogeneities in phase transitions that occur in the early universe as it expands and the laws of fundamental physics undergo spontaneous symmetry breaking (Dolgov et al., 2002; Hogan, 1986; Kosowsky et al., 2002; Witten, 1984). The epoch that is of most interest to LISA is when the temperature of the universe was around 1 TeV. At this epoch the particle horizon size was equal to the gravitational wavelength in LISA's band, blueshifted by the change in the size of the universe back to that time. Coincidentally, this temperature corresponds to the temperature at which the electroweak symmetry breaking transition occurred. Current models of this transition favor its being a second-order transition, without strong density inhomogeneities, but the situation could well be different due to effects outside the Standard Model of the fundamental interactions.

Cosmic strings have been suggested as sources of gravitational waves in the LISA band (Damour and Vilenkin, 2005), contributing to a stochastic background signal (Depies and Hogan, 2007; Hogan, 2006a).

Finally there have been many suggestions of alternative models of inflation, of the universe before inflation, and of brane-world alternatives to inflation that lead to different predictions of a gravitational-wave background (Hogan, 2000; Randall and Servant, 2007). Some models contain parameters that could lead to a spectrum peaked in the LISA band (Buonanno et al., 1997; Easther and Lim, 2006; Felder and Kofman, 2007), with little radiation at ground-based frequencies and at the low frequencies observed in the CMB.

With so many potential backgrounds, it is important to ask how LISA might distinguish one from another. The Galactic background is the most distinctive, being anisotropic. By using the angular dependence of the measured background at low frequencies along with extrapolations from the measured compact-binary population at high frequencies it should be possible to estimate how much of the measured background is Galactic. Extragalactic backgrounds are harder to separate. The extragalactic binary background is expected to have a flatter frequency dependence (in amplitude h) (Farmer and Phinney, 2003) than the scale-free background predicted by inflation, up to a cutoff determined by the minimum orbital period around a white dwarf. The extragalactic EMRI background would extend up to higher frequencies than this and would be predicted by the statistics of the observed resolved EMRI events. Backgrounds due to phase transitions or exotic pre-inflation physics could show peaks in the LISA band that would not be expected from binary populations (Buonanno et al., 1997; Hogan, 1986). If a background is detected at frequencies above 1 mHz, it is inevitable that there will be much modeling and discussion before confident identifications are made. However, the observation itself will be one of the most important LISA could make.

6.2. Science return from detecting these sources

The detection of a cosmological background from the early universe would be an observation of the greatest significance for fundamental physics and astronomy. Gravitational waves, and possibly neutrinos, are likely to be the only way we will get direct information about epochs earlier than decoupling ($z \sim 1100$). Seeing phase transitions, pre-inflation physics, or inflation itself provide welcome experimental input to the search for models of fundamental physics at energies well above those that can be explored in ground-based accelerator physics.

The detection of the Galactic background is expected: if it is not seen then either LISA is not operating correctly or there is something very wrong with our understanding of binaries in the Galaxy. However, the spectrum, degree of anisotropy, and population of resolved sources at the edge of the confusion limit will all contain useful astrophysical information about the population of compact-object binaries.

The detection of a background from extragalactic compact binaries would similarly be input to models of compact objects and their evolution in external galaxies, would inform us whether our Galaxy is similar to or different from others in this respect, and would contain information about the long-term evolution of binary systems. If the background comes from EMRIs then that will contain information about the stellar populations in the centers of galaxies and about the population statistics of central black holes.

6.3. Characteristics of the signals

Ideal stochastic signals are by definition fields of normally-distributed, randomly fluctuating gravitational-wave amplitudes, shaped by a spectrum and possibly by an anisotropic distribution of arrival directions. At any given frequency the wave field is partly correlated between separated points, the correlations falling off on distance scales of the wavelength. This correlation allows separated detectors to detect signals by cross-correlating their outputs. Gravitational waves from the early universe are well described by this ideal model.

Waves from numerous individual astrophysical sources satisfy this model if the central limit theorem applies: that is, if there are a sufficient number of superposed independent sources. The background due to extragalactic binaries is expected to satisfy this model well (Farmer and Phinney, 2003). The Galactic binary background at low frequencies and any confusion noise due to EMRI sources satisfies this model only approximately, however, corrections to the detection threshold in matched-filter searches are small (Racine and Cutler, 2007). Both source populations contain very numerous emitters at large distances, but they also contain more sparsely distributed nearby sources, so that the statistics of the fluctuating field are only approximately Gaussian. This noise is called *cyclo-stationary*. At the very edge of the noise the statistics get even more unpredictable.

6.4. Goals of data analysis

LISA should detect and measure the spectral content and the degree of anisotropy of the Galactic binary stochastic background at frequencies below about 1 mHz. At higher frequencies LISA should have the capability of detecting or setting limits on random Gaussian wave fields at all frequencies in its band. If it detects a wave field it should measure or limit the degree of anisotropy.

In addition LISA should resolve and measure as many discrete sources at the edge of the stochastic distributions as possible.

6.5. Requirements for data analysis

The principal requirement for data analysis is to distinguish gravitational-wave noise from instrumental noise.

The method available to ground-based detectors, cross-correlation of independent data streams (Allen and Romano, 1999), is not helpful for LISA when dealing with isotropic signals (Kudoh and Taruya, 2005). It is possible to construct two different detectors from any pair of LISA's arms and to cross-correlate their output, but they always share a common arm and presumably therefore share roughly half their noise in common; it turns out that this is just enough common noise to prevent them from gaining any advantage from cross-correlation.

Instead, LISA has two ways of directly estimating its instrumental noise, so that if the observed noisy signal is larger than, of even comparable to the estimate of the instrumental noise then LISA can claim a detection.

1. The first way is simply to understand the instrument, using as much monitoring of auxiliary data channels as possible to reconstruct a reliable instrumental noise model.

2. The second way is more interesting and more reliable: at low frequencies there is a certain combination of the TDI variables, called the *symmetrized Sagnac mode* (Armstrong et al., 1999; Sylvestre and Tinto, 2003), which is insensitive to gravitational waves. The signal in this channel will be a particular linear combination of the noise in the different TDI channels, and so its power will presumably limit the power in the instrumental noise in all TDI channels (Hogan and Bender, 2001; Tinto et al., 2001). If the noise power in the TDI channels that do respond to gravitational-wave signals is larger than this, then this must come from gravitational waves. This works best at frequencies below 1 mHz, where the gravitational wavelength is large compared to the LISA arm length. Above a few mHz this second method will not help to identify backgrounds.

LISA can search for anisotropy by two methods. One method takes advantage of the different antenna patterns arising from different TDI combinations and analyzes correlations between the data streams generated by different TDI variables (Seto and Cooray, 2004; Taruya, 2006; Taruya and Kudoh, 2005); the other method relies on the time-varying strength of the confusion foreground as the antenna pattern due to a single TDI variable sweeps about the sky (Cornish, 2001; Edlund et al., 2005b; Ungarelli and Vecchio, 2001). In either approach, the important observable is the variation of the stochastic foreground as the quadrupolar antenna pattern of a TDI variable sweeps past the Galaxy. Techniques developed for ground based observations to set limits on the anisotropy are also rapidly maturing and amenable to be adapted in a sufficiently straightforward way to the LISA case (Allen and Ottewill, 1997; Ballmer, 2006).

Data analysis must take into account that there are likely to be non-Gaussian artifacts in both the instrumental noise (as it is seen in ground-based detectors) and in the gravitational-wave signal. The latter are caused by residuals from the removal of resolved sources, and also near the edge of the confusion noise, where there are few overlapping sources. Careful planning is required for data analysis to take account of these corrections.

6.6. Progress in development of data-analysis methods

In the last few years, work has progressed from designing data-analysis algorithms that appear in principle to be able to attain the required goals to the application of theses algorithms in the Mock LISA Data Challenge. Stochastic background signals have been included first in the second round of the MLDC and more fully in the third round of the MLDC (section 8.2.4)

Experience with ground-based stochastic searches suggests that the principal problem in implementing algorithms is dealing with instrumental artifacts; however very considerable progress has been made in the past years, in particular by tackling the joint analysis of the data sets from the two Hanford interferometers that share the same vacuum (Lazzarini et al., 2004) which contain a significant degree of correlated noise and as such provide an excellent testbed for techniques applicable to LISA. The search for a stochastic background in the S3 and S4 runs of LIGO (Abbott et al., 2005; Abbott et al., 2007) is another example for an appplication of algorithms designed to find stochastic signals.

For LISA there will be, as mentioned above, the problem of dealing with residuals of strong signals. These issues will only be addressed once the global search method for resolvable sources has been fully understood.

6.7. Summary and conclusions

The stochastic background of gravitational radiation is one of the most important signals LISA could detect, and LISA has unique ways of independently measuring instrumental noise in order to distinguish it from a gravitational-wave background noise. At low frequencies LISA must measure a cyclostationary noise consisting of superposed signals from large numbers of compact-object binaries in the Galaxy. If the early universe produced a background at a level of 10^{-10} of the cosmological closure density in the LISA band, LISA has an excellent chance of detecting it. This would be a discovery of fundamental importance in physics. There seem to be no serious obstacles to performing these measurements with LISA.

6.8. References

B. Abbott, R. Abbott, R. Adhikari, J. Agresti, P. Ajith, et al. (LIGO Scientific Collaboration). Upper Limits on a Stochastic Background of Gravitational Waves. *Phys. Rev. Lett.*, 95(22):221101, November 2005, also as arXiv:astro-ph/0507254.

B. Abbott, R. Abbott, R. Adhikari, J. Agresti, P. Ajith, et al. Searching for a Stochastic Background of Gravitational Waves with the Laser Interferometer Gravitational-Wave Observatory. *ApJ*, 659:918–930, April 2007, also as arXiv:astro-ph/0608606.

B. Allen and A. C. Ottewill. Detection of anisotropies in the gravitational-wave stochastic background. *Phys. Rev. D*, 56:545–563, July 1997, also as arXiv:gr-qc/9607068.

B. Allen and J. D. Romano. Detecting a stochastic background of gravitational radiation: Signal processing strategies and sensitivities. *Phys. Rev. D*, 59(10): 102001–+, May 1999, also as arXiv:gr-qc/9710117.

J. W. Armstrong, F. B. Estabrook, and M. Tinto. Time-Delay Interferometry for Space-based Gravitational Wave Searches. *ApJ*, 527:814–826, December 1999.

S. W. Ballmer. A radiometer for stochastic gravitational waves. *Class. Quantum Grav.*, 23:179–+, April 2006, also as arXiv:gr-qc/0510096.

L. Barack and C. Cutler. Confusion noise from LISA capture sources. *Phys. Rev. D*, 70:122002–+, December 2004, also as arXiv:gr-qc/0409010.

P. L. Bender and D. Hils. Confusion noise level due to galactic and extragalactic binaries. *Class. Quantum Grav.*, 14:1439–1444, June 1997.

A. Buonanno, M. Maggiore, and C. Ungarelli. Spectrum of relic gravitational waves in string cosmology. *Phys. Rev. D*, 55:3330–3336, March 1997, also as arXiv:gr-qc/9605072.

S. Chongchitnan and G. Efstathiou. Prospects for direct detection of primordial gravitational waves. *Phys. Rev. D*, 73(8):083511–+, April 2006, also as arXiv:astro-ph/0602594.

N. J. Cornish. Mapping the gravitational-wave background. *Class. Quantum Grav.*, 18:4277–4291, October 2001, also as arXiv:astro-ph/0105374.

T. Damour and A. Vilenkin. Gravitational radiation from cosmic (super)strings: Bursts, stochastic background, and observational windows. *Phys. Rev. D*, 71(6): 063510–+, March 2005, also as arXiv:hep-th/0410222.

M. R. Depies and C. J. Hogan. Stochastic gravitational wave background from light cosmic strings. *Phys. Rev. D*, 75(12):125006–+, June 2007, also as arXiv:astro-ph/0702335.

A. D. Dolgov, D. Grasso, and A. Nicolis. Relic backgrounds of gravitational waves from cosmic turbulence. *Phys. Rev. D*, 66(10):103505–+, November 2002, also as arXiv:astro-ph/0206461.

R. Easther and E. A. Lim. Stochastic gravitational wave production after inflation. *Journal of Cosmology and Astro-Particle Physics*, 4:10–+, April 2006, also as arXiv:astro-ph/0601617.

J. A. Edlund, M. Tinto, A. Królak, and G. Nelemans. Simulation of the white dwarf white dwarf galactic background in the LISA data. *Class. Quantum Grav.*, 22: S913–S926, September 21 2005a, also as arXiv:gr-qc/0504026.

J. A. Edlund, M. Tinto, A. Królak, and G. Nelemans. White-dwarf white-dwarf galactic background in the LISA data. *Phys. Rev. D*, 71:024015–+, June 2005b, also as arXiv:gr-qc/0504112.

A. J. Farmer and E. S. Phinney. The gravitational wave background from cosmological compact binaries. *MNRAS*, 346:1197–1214, December 2003, also as arXiv:astro-ph/0304393.

G. N. Felder and L. Kofman. Nonlinear Inflaton Fragmentation after Preheating. *Phys. Rev. D*, 75(4):043518–+, February 2007, also as arXiv:hep-ph/0606256.

A. H. Guth. Inflationary universe: A possible solution to the horizon and flatness problems. *Phys. Rev. D*, 23(2):347–356, Jan 1981.

D. Hils and P. L. Bender. Gravitational Radiation from Helium Cataclysmics. *ApJ*, 537:334–341, July 2000.

D. Hils, P. L. Bender, and R. F. Webbink. Gravitational radiation from the Galaxy. *ApJ*, 360:75–94, September 1990.

C. J. Hogan. Gravitational radiation from cosmological phase transitions. *MNRAS*, 218:629–636, February 1986.

C. J. Hogan. Gravitational waves from light cosmic strings: Backgrounds and bursts with large loops. *Phys. Rev. D*, 74(4):043526–+, August 2006a, also as arXiv:astro-ph/0605567.

C. J. Hogan. Gravitational Wave Sources from New Physics. In *AIP Conf. Proc. 873: Laser Interferometer Space Antenna: 6th International LISA Symposium*, pages 30–40, November 2006b, also as arXiv:astro-ph/0608567.

C. J. Hogan. Gravitational Waves from Mesoscopic Dynamics of the Extra Dimensions. *Phys. Rev. Lett.*, 85:2044–2047, September 2000, also as arXiv:astro-ph/0005044.

C. J. Hogan and P. L. Bender. Estimating stochastic gravitational wave backgrounds with the Sagnac calibration. *Phys. Rev. D*, 64(6):062002–+, September 2001, also as arXiv:astro-ph/0104266.

A. Kosowsky, A. Mack, and T. Kahniashvili. Gravitational radiation from cosmological turbulence. *Phys. Rev. D*, 66(2):024030–+, July 2002, also as arXiv:astro-ph/0111483.

H. Kudoh and A. Taruya. Probing anisotropies of gravitational-wave backgrounds with a space-based interferometer: Geometric properties of antenna patterns and their angular power. *Phys. Rev. D*, 71(2):024025–+, January 2005, also as arXiv:gr-qc/0411017.

A. Lazzarini, S. Bose, P. Fritschel, M. McHugh, T. Regimbau, et al. Optimal combination of signals from colocated gravitational wave interferometers for use in searches for a stochastic background. *Phys. Rev. D*, 70(6):062001, September 2004, also as arXiv:gr-qc/0403093.

A. D. Linde. Scalar field fluctuations in the expanding universe and the new inflationary universe scenario. *Physics Letters B*, 116:335–339, October 1982.

M. Maggiore. Gravitational wave experiments and early universe cosmology. *Phys. Rep.*, 331:283–367, 2000, also as arXiv:gr-qc/9909001.

G. Nelemans, L. R. Yungelson, and S. F. Portegies Zwart. The gravitational wave signal from the Galactic disk population of binaries containing two compact objects. *A&A*, 375:890–898, September 2001, also as arXiv:astro-ph/0105221.

L. R. Price and X. Siemens. Stochastic backgrounds of gravitational waves from cosmological sources: Techniques and applications to preheating. *Phys. Rev. D*, 78(6):063541–+, September 2008, also as arXiv:0805.3570.

É. Racine and C. Cutler. Gaussianity of LISA's confusion backgrounds. *Phys. Rev. D*, 76(12):124033–+, December 2007, also as arXiv:0708.4242.

L. Randall and G. Servant. Gravitational Waves from Warped Spacetime. *Journal of High Energy Physics*, 5:54–+, May 2007, also as arXiv:hep-ph/0607158.

N. Seto and A. Cooray. LISA measurement of gravitational wave background anisotropy: Hexadecapole moment via a correlation analysis. *Phys. Rev. D*, 70: 123005–+, December 2004, also as arXiv:astro-ph/0403259.

G. Sigl, J. Schnittman, and A. Buonanno. Gravitational-wave background from compact objects embedded in active galactic nuclei accretion disks. *Phys. Rev. D*, 75(2):024034–+, January 2007, also as arXiv:astro-ph/0610680.

D. N. Spergel, R. Bean, O. Dore', M. R. Nolta, C. L. Bennett, et al. Wilkinson Microwave Anisotropy Probe (WMAP) Three Year Results: Implications for Cosmology. *ApJS*, 170:377–408, June 2007, also as arXiv:astro-ph/0603449.

A. A. Starobinskiǐ. Spectrum of relict gravitational radiation and the early state of the universe. *Journal of Experimental and Theoretical Physics Letters*, 30:682–+, December 1979.

J. Sylvestre and M. Tinto. Noise characterization for LISA. *Phys. Rev. D*, 68:102002–+, November 2003, also as arXiv:gr-qc/0308085.

A. Taruya. Probing anisotropies of gravitational-wave backgrounds with a space-based interferometer. III. Reconstruction of a high-frequency sky map. *Phys. Rev. D*, 74(10):104022, November 2006, also as arXiv:gr-qc/0607080.

A. Taruya and H. Kudoh. Probing anisotropies of gravitational-wave backgrounds with a space-based interferometer. II. Perturbative reconstruction of a low-frequency skymap. *Phys. Rev. D*, 72(10):104015–+, November 2005, also as arXiv:gr-qc/0507114.

The Planck Consortium. Planck – The Scientific Programme, 2005. online at `http://www.tapir.caltech.edu/dokuwiki/listwg1b:home`.

M. Tinto, J. W. Armstrong, and F. B. Estabrook. Discriminating a gravitational wave background from instrumental noise in the LISA detector. *Phys. Rev. D*, 63: 4081–4086, January 15 2001.

C. Ungarelli and A. Vecchio. Studying the anisotropy of the gravitational wave stochastic background with LISA. *Phys. Rev. D*, 64:121501–+, December 15 2001, also as arXiv:astro-ph/0106538.

E. Witten. Cosmic separation of phases. *Phys. Rev. D*, 30:272–285, July 1984.

7. Unexpected Signals: LISA's Discovery Space

Each time astronomers have opened a new observational window the universe has provided big surprises. The cosmic microwave background radiation, pulsars, X-ray binaries (with their black holes), gamma-ray bursts, and many other phenomena were discovered serendipitously with instruments built with no prior expectation of observing them. It is therefore very likely that the gravitational-wave universe will offer big surprises. LISA, observing the low-frequency band for the first time, and operating with unprecedented sensitivity, may well be the instrument that makes the first unexpected discoveries. Any such discoveries would be likely to make a large impact on our understanding of the universe.

Finding the unexpected presents unique problems in gravitational wave detection because, as explained earlier, the basis of data analysis is filtering for signals using previously computed templates. Unexpected sources, by definition, do not have templates. Special methods therefore have to be designed and implemented for this problem.

Searches for unexpected signals are a standard and major part of the data-analysis activity of the ground-based gravitational-wave detector projects. The LIGO Scientific Collaboration (LSC) operates four data-analysis teams looking for different kinds of sources, and the "burst group" that looks for unexpected and unmodeled signals is the largest. Therefore, LISA can draw on considerable legacy experience from the ground-based projects.

7.1. Description of the sources

Many predictions have been made for exotic sources of gravitational waves that are not part of the principal science case for LISA (see Hogan, 2006). These include cosmic string bursts (Damour and Vilenkin, 2005; Depies and Hogan, 2007), burst emitted by close encounters of compact objects with MBH (Rubbo et al., 2006; Yunes et al., 2008) and gravitational collapse to form massive black holes. Cosmic strings arise naturally during the expansion of the universe in some scenarios for spontaneous symmetry breaking in fundamental physics. Cosmic string loops generically develop "cusps" of order once per oscillation. On theses cusps, a point of the string is accelerated to the speed of light for an instant and it beams a burts of GW along its direction of motion. Gravitational collapse of massive gas clouds to black holes in the range $10^3\,M_\odot$ to $10^6\,M_\odot$ is possible (Begelman et al., 2006; Loeb and Rasio, 1994, see for instance), although it is not the preferred method for forming black holes in this mass range: most theorists seem to prefer models in which black holes grow from smaller sizes either by gas accretion or by merging with other holes. Signals from collapse are hard to predict because they depend on the details of the initial state of the collapsing gas cloud.

The collapse of a very massive star ($10^3\,M_\odot$) to a black hole can lead to a detectable gravitational wave when the star's magnetic field and the remaining matter interact (Liu et al., 2007).

Of course, it may be that there are sources we do not anticipate at all. This might arise in scenarios where dark matter contains a minority component that forms compact objects (*e. g.*, massive interacting boson fields), or in braneworld scenarios where matter on a nearby brane is able to exert gravitational effects on our brane but cannot interact with us electromagnetically. These possibilities are speculative but cannot be dismissed, principally because our understanding of the fundamental interactions beyond the Standard Model is poor. Detections of these kinds would, therefore, have extraordinary importance for fundamental physics.

7.2. Science return from detecting these sources

If cosmic strings are seen by LISA, this will partially confirm and also constrain string-theory models of the fundamental interactions. If the inferred density of strings is great enough, there will be implications for structure formation in the early universe.

Detecting bursts from massive gravitational collapse would provide unique insight into the conditions of gas in the early universe and the pathway for the formation of the massive black holes that appear to be ubiquitous in galaxies.

Detections of signals from dark matter or from nearby branes would be much harder to interpret. If the detections are secure then they will stimulate a great deal of theoretical work to try to discover the sources. Subsequent events will be easier to recognize and will undoubtedly provide important further information for their interpretation. Models may also suggest follow-up detection channels, such as ground-based gravitational-wave observations, or the use of gravitational lensing. The results of this work would undoubtedly be revolutionary for fundamental physics.

7.3. Characteristics of the signals

Although detailed waveforms will not be known for this class of sources, there may be limited information about some of them. Cosmic string bursts, for example, have clear predictions about the time-dependence of the signal amplitude as it rises and then falls again because the beam sweeps past LISA. Bursts from gravitational collapse are expected to be broadband events with durations comparable to the light-crossing time over the black hole that is formed; the black-hole ringdown after the event would be a convincing signature of the formation of a black hole.

In general, one can enhance sensitivity by looking for signals with particular properties. It is plausible that any system consisting of massive compact objects (say boson stars in the dark matter) would radiate most strongly if it is dominated by rotation (*e. g.*, orbits or spins). Then the signal would be relatively narrow-band around some central frequency, or around a frequency that changes slowly. Coalescing boson stars would produce a signal similar to that of coalescing black holes until tidal interactions or the merger cuts it off. They might show up in searches for black hole mergers, and a targeted search would be easy to construct by modifying existing black-hole–merger search codes.

7.4. Goals of data analysis

Data analysis will be designed to reveal unexpected signals even when they compete with stronger signals for which template families exist. Several different kinds of signals should be allowed for: exotic but predicted signals, generic narrow-band signals, and possibly others.

7.5. Requirements for data analysis

The requirement that it be possible to recognize unexpected signals even in the presence of confusion from signals that belong to expected template families is a stringent one. Conceptually it is easiest to think of this as a search that is performed on a "cleaned" data stream, produced by subtracting the identified signals with known waveforms. In practice this might of course be done as part of a global fit. Any global fitting method should therefore demonstrate that it leaves such unexpected signals untouched when it identifies other signals for which it has templates. The analysis should aim at recognizing unexpected signals if they have SNR bigger than 10. This places constraints on the residuals left from the identification of stronger signals.

Data analysis must take into account that LISA contains information of a "coincidence" nature: the different TDI data streams must register the event in a consistent way. One way to test this is to construct the *null stream* or *zero signal* combination of TDI variables, which is a combination that eliminates all gravitational-wave signals coming from a particular location on the sky (Tinto and Larson, 2004). If that location is the one that a tentatively detected signal comes from, and if the signal is still present in the null stream, then the signal is not real. This can be a very effective veto against instrumental artifacts. This veto is already in use in the ground-based searches for gravitational waves, where the output from three or more separated detectors plays the role that TDI data streams do in LISA.

7.6. Progress in development of data-analysis methods

Ground-based searches have already generated considerable experience with generic data-analysis methods for unexpected sources. Analysts working on these searches have implemented *null stream* vetos (Wen and Schutz, 2005); they have time-frequency and other methods for doing searches (Abbott et al., 2007, 2008); detector networks are used to deal with non-gaussian noise (Principe and Pinto, 2009); they are presently implementing a search for cosmic string signals; and they are developing databases of test signals that embody various hypotheses about the general properties of such signals. There has so far not been very much work on this problem specifically in the LISA context. It would probably not be wise to do this until the global search method has been chosen. Indeed, one criterion for choosing the global search method must be that it leave unexpected signals undisturbed, which can be tested using simulated test signals, as is done in ground-based analysis. The next step after that should be to add on top of the global solution a search for unexpected signals in the residuals.

7.7. Summary and conclusions

Searching for events in LISA's "discovery space" will be one of the most challenging and potentially important analysis steps for the LISA data. Once global searches for expected signal families are available, the development of searches for remaining unexplained and unexpected signals must be a high priority. Such searches need to be kept up to date, because theoretical computations and astronomical discoveries constantly supply more information that helps to tune and target these searches.

7.8. References

B. Abbott, R. Abbott, R. Adhikari, J. Agresti, P. Ajith, et al. Search for gravitational-wave bursts in LIGO data from the fourth science run. *Class. Quantum Grav.*, 24: 5343–5369, November 2007, also as arXiv:0704.0943.

B. Abbott, R. Abbott, R. Adhikari, P. Ajith, B. Allen, et al. Search for Gravitational-Wave Bursts from Soft Gamma Repeaters. *Physical Review Letters*, 101(21):211102–+, November 2008, also as arXiv:0808.2050.

M. C. Begelman, M. Volonteri, and M. J. Rees. Formation of supermassive black holes by direct collapse in pre-galactic haloes. *MNRAS*, 370:289–298, July 2006, also as arXiv:astro-ph/0602363.

T. Damour and A. Vilenkin. Gravitational radiation from cosmic (super)strings: Bursts, stochastic background, and observational windows. *Phys. Rev. D*, 71(6): 063510–+, March 2005, also as arXiv:hep-th/0410222.

M. R. Depies and C. J. Hogan. Stochastic gravitational wave background from light cosmic strings. *Phys. Rev. D*, 75(12):125006–+, June 2007, also as arXiv:astro-ph/0702335.

C. J. Hogan. Gravitational Wave Sources from New Physics. In *AIP Conf. Proc. 873: Laser Interferometer Space Antenna: 6th International LISA Symposium*, pages 30–40, November 2006, also as arXiv:astro-ph/0608567.

Y. T. Liu, S. L. Shapiro, and B. C. Stephens. Magnetorotational collapse of very massive stars to black holes in full general relativity. *Phys. Rev. D*, 76(8):084017–+, October 2007, also as arXiv:0706.2360.

A. Loeb and F. A. Rasio. Collapse of primordial gas clouds and the formation of quasar black holes. *ApJ*, 432:52–61, September 1994, also as arXiv:astro-ph/9401026.

M. Principe and I. M. Pinto. Locally optimum network detection of unmodelled gravitational wave bursts in an impulsive noise background. *Class. Quantum Grav.*, 26(4):045003–+, February 2009.

L. J. Rubbo, K. Holley-Bockelmann, and L. S. Finn. Event Rate for Extreme Mass Ratio Burst Signals in the Laser Interferometer Space Antenna Band. *ApJ*, 649: L25–L28, September 2006.

M. Tinto and S. L. Larson. LISA time-delay interferometry zero-signal solution: Geometrical properties. *Phys. Rev. D*, 70:062002, September 15 2004.

L. Wen and B. F. Schutz. Coherent network detection of gravitational waves: the redundancy veto. *Class. Quantum Grav.*, 22:1321–+, September 2005, also as arXiv:gr-qc/0508042.

N. Yunes, C. F. Sopuerta, L. J. Rubbo, and K. Holley-Bockelmann. Relativistic Effects in Extreme Mass Ratio Gravitational Wave Bursts. *ApJ*, 675:604–613, March 2008, also as arXiv:0704.2612.

8. Collective activities

8.1. Introduction

Significant progress has been made in the recent past on a variety of open problems related to data analysis and waveform computations that are essential for the success of the mission. Open problems still remain that require a coordinated effort with clearly identified priorities to ensure the maximum science return from the mission. Here we provide details on the future program.

8.2. Mock LISA Data Challenge

The work of the international community to tackle open data analysis issues steadily intensified during the past few years, as is documented in the previous sections of the document. It was recognized early on that there was a great need for overall coordination and organization of this effort to insure that all necessary know-how for fully exploiting the missions scientific data be available at the proper time. In consequence, the Working Group on Data Analysis (LIST-WG1B) decided at the LISA International Science Team (LIST) meeting of December 2005 to organize several rounds of Mock LISA Data Challenges (MLDC, with the dual purpose of (i) fostering the development of LISA data analysis tools and capabilities, and (ii) demonstrating the technical readiness already achieved by the gravitational-wave community in distilling a rich science payoff from the LISA data output. The Mock LISA Data Challenges were also proposed and discussed at meetings organized by the US and European LISA Project that were attended by a broad cross section of the international GW community. These challenges are meant to be blind tests, but not competitions; the greatest scientific benefit gained will have little to do with who is the most successful contestant, but will instead be derived from comparing in a quantitative way the different results, analysis methods, and implementations.

A Mock LISA Data Challenge (MLDC) Task Force was constituted at the beginning of 2006 and has been working ever since to formulate challenge problems of maximum efficacy, to establish criteria for the evaluation of the analyzes, to develop standard models of the LISA mission and GW sources, to provide computing tools—LISA response simulators, source waveform generators, and a Mock Data Challenge file format—and more generally to provide any technical support to the challengers which might prove necessary. The challenges involve the distribution of several data sets, encoded in a simple standard format, and containing combinations of realistic simulated LISA noise with the signals from one or more GW sources of parameters unknown to the challenge participants, who are asked to return the maximum amount of correct information about the sources, and to produce technical notes detailing their work.

In order to ensure the incremental development of data analysis approaches, software, pipelines and infrastructure, the Task Force has decided to issue several rounds of challenges, at intervals of approximately 6 months, that involve progressively more complex GW signals and noise realizations. Each round of MLDCs consists of multiple data sets with signals of different nature and strength embedded in synthetic noise. Two classes of data sets are distributed at each release: the proper challenge data sets (for which the source parameters are unknown) and training data sets, with GW signals of similar nature to those included in the blind tests but whose parameters are made public (for validation purposes, the training data sets are distributed in two flavors: "noise-free" and "noisy").

All the software and the necessary documentation to generate data sets are public (MLDC webpage) and open source, so that interested parties can produce their own mock data streams. This approach ahs been chosen to facilitate the development of analysis tools and the testing and validation of the algorithms. The challenge data sets are generated by two members of the Task Force who do *not* take part in the challenges and are the only repositories of the key to recover the source parameters. Progress on the work is monitored through extended teleconferences. MLDC results, including analyzes of how well different methods performed, will be disseminated through technical documents and articles in peer reviewed journals.

So far, the MLDC consists of three distinct rounds and a re-run of the first run.

8.2.1. Challenge 1

The first round of challenges was aimed at enabling the development of the necessary data-analysis building blocks for simple LISA data analysis tasks. It focused on data sets that contain a single source (generally with moderate-to-high SNR) or a small number of sources that do not overlap in the frequency domain; an important exception was however made for two data sets that contain a few tens of galactic binary systems whose signals overlap significantly in a small frequency region. These latter data sets provided a fairly realistic representation of the novel aspect in GW data analysis offered by the LISA mission, though restricted to a limited frequency band and a particular signal class. The first round of challenges concentrated on sources currently listed as minimum science requirements (LISA International Science Team, 2006): galactic binaries, including verification binaries, and massive black hole binary systems. Data sets containing EMRIs are also made available in the first round, in order to facilitate the development of data-analysis tools for this particular demanding class of sources. However, the results of the challenge were due only in June 2007, together with Challenge 2.

All the data sets, with the exception of the EMRI challenge data discussed above, represent an observation period of one year and contain instrumental noise modeled as Gaussian and stationary with the additional assumption that the laser frequency noise has been exactly removed. No foreground radiation was included. More details are provided in Arnaud et al. (2006); MLDC omnibus; MLDC webpage.

Results

Ten groups took part in the first round of MLDCs and submitted results by the deadline of 4th December 2006.[1] Results from each group—including a technical note describing the analysis method, its implementation and the results, and files with the "best fit parameters"—are posted on the MLDC website (MLDC webpage). An overview and summary of the first round has been published by the MLDC Task Force (Arnaud et al., 2007a)

Each Challenge data set, including the two data sets containing an unknown number of overlapping galactic binary signals was analysed by at least two groups; several data sets were tackled by three or more collaborations for a total of seven groups working on the galactic binaries challenge data sets and three groups devoting attention to the massive black hole binary challenge data sets, pursuing a number of different analysis techniques.

Challenge data sets were tackled with template bank based methods, Markov Chain Monte Carlo (MCMC) methods, Hilbert transform, tomographic reconstruction, and time-frequency methods. In some cases, different techniques were used in combination to produce a multi-stage or hierarchical analysis. Template bank based methods and MCMC were applied to both galactic binaries data sets and massive–black-hole data sets. The entries showed a large disparity in maturity of development of analysis pipelines, as was to be expected at the beginning of the development work.

The results submitted by the groups varied in quality – most of the groups clearly detail in their report why the results should just be considered as preliminary – but, collectively, *all* the signals were recovered correctly, according to figures of merit described in the Task Force summary report, see (MLDC Wiki). Due to the different stage of development of the analysis approaches, a strict comparison of the performance of the algorithms would have been premature, and therefore was not being carried out.

[1] The MLDC Task Force is aware of a handful of other groups who have been developing analysis algorithms to tackle the first round of MLDCs, but results were not produced in time for submission. One can therefore be moderately optimistic about an even more numerous participation to the second round of MLDCs.

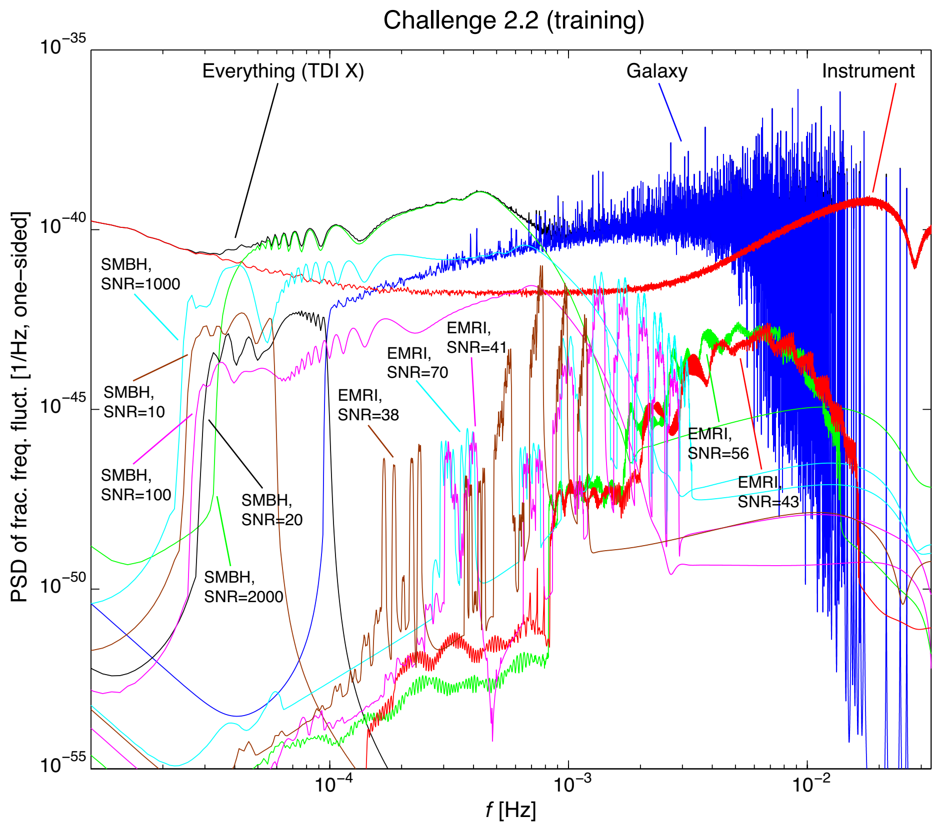


Figure 8.1.: Power spectral density of all the simulated signals that appear in the training data for the second round of MLDC. The simulated LISA data streams for this realization contain 4 SMBH, 5 EMRI, and 26.1 Million Galactic Binaries.

8.2.2. Challenge 2

Challenge 2 (Arnaud et al., 2007b), issued in January 2007 with results due at the end of June 2007, raised the bar by proposing three complex subchallenges. Data set 2.1 contained a full population of Galactic binary systems (about 26 million sources). Data set 2.2 contained a different realization of the Galaxy, plus 4–6 MBH binary inspirals with single-interferometer signal-to-noise ratios (SNRs) between 10 and 2000 and a variety of coalescence times, and five EMRI with an SNR between 30 and 100. Last, five more data sets (denoted 1.3.1–5, since they were actually released at the time of Challenge 1) contained a single EMRI signal over instrument noise alone.

Results

The results of the second round are discussed in detail in the report of the MLDC task force (Babak et al., 2008a) on which this subection is based. All the solutions submitted by participating groups, together with technical write-ups of their methods and findings, can be found at the URL www.tapir.caltech.edu/~mldc/results2.

Thirteen collaborations submitted a total of 22 entries, including a proof-of-principle analysis for stochastic backgrounds performed on data set 2.1. Altogether, Challenge 2 successfully demonstrated the identification of $\sim 20,000$ Galactic binaries, the accurate estimation of MBH inspiral parameters, and the positive detection of EMRIs.

Five groups submitted Galactic-binary catalogs for data sets 2.1 and 2.2 using a variety of methods, such as Reverse Jump MCMC, $\mathcal{F}$-statistics and MHMC algorithms.

Although an assessment of the success of these methods is difficult, as the precise number of "detectable" sources is unknown and the notion of "false positives" is somewhat ill-defined, these challenges demonstrated a solid capability in analyzing signals from the Galaxy and resolving a large number of binaries.

Four groups reported parameter sets for the MBH binaries in data set 2.2, employing statistical methods such as MHMC, hybrid methods that combined $\mathcal{F}$-statistics with template-bank mathced filtering and MCMC, and time-frequency based methods. Most of the groups succeeded in positively detecting the four MBH binaries in the data set.

Altogether, this challenge demonstrated a solid capability in the detection and parameter estimation of MBH inspirals with moderate SNR, even in the presence of a strong Galactic background, at least if the inspirals can be considered close to the idealized model: circular and adiabatic with negligible spin effects.

Three groups reported parameter sets for the EMRIs in data sets 1.3.1–1.3.4; the problem of detecting these systems on top of the Galactic background was not addressed by any group.

Challenge 2 has given a convincing demonstration that a significant portion of the LISA science objectives can be achieved with techniques that are currently in hand. Most of the research groups that participated in Challenge 1 have successfully made the transition to the greater complexity of Challenge 2.

8.2.3. Challenge 1B

The very steep increase in complexity introduced over a short time-scale with Challenge 2 and the need to consolidate analysis techniques (especially so for EMRIs) before moving to even more taxing challenges motivated the organization of Challenge 1B, a repeat of Challenge 1 with the addition of single-EMRI data sets. Challenge-1B data sets were distributed in the late summer 2007, with a deadline of December 2007 for entries. Ten collaborations submitted solutions. Highlights from this round include the range of techniques used, the participation of a number of new groups that successfully recovered signals from galactic binaries and MBH binaries, and the first convincing demonstration of EMRI detection *and* parameter estimation.

Results

A detailed account of the results of Challenge 1B can be found in (Babak et al., 2008b), technical notes accompanying the entries to Challenge 1B can be found at www.tapir.caltech.edu/~mldc/results1B/results.html.

Five groups submitted entries for the monochromatic galactic binaries, employing a range of techniques similar to those used in previous challenges. New implementations and different technical solutions are being pursued, such as using code developed for other purposes (XSPEC, an X-Ray spectral fitting package (XSP)) or genetic algorithms to maximize the $\mathcal{F}$-statistic.

Detection of the 25 "verification binaries" was undertaken by three groups with good results. All the participating gropus failed to detect the 20 random binaries. Here, a bug in the generation of the source paramaters resulted in a too low SNR for the binaries, so that the failure to detect any of them is consistent.

Altogether, Challenge 2 provided a more forceful demonstration of LISA's science objectives for Galactic binaries; but Challenge 1B was still very useful for new groups to start implementing search methods, and for established groups to continue tuning them.

Two groups submitted entries for the MBH, using different techniques such as a a three-step hierarchical strategy combining a time–frequency track-search analysis, a template-bank matched-filtering search, and a final MCMC stage to evaluate the posterior probability densities of source parameters and a stochastic–template-bank matched-filtering search (Harry et al., 2008).

Both groups succeeded in making very accurate predictions of the time of coalescence. However, the retreived sky positions varied significantly due to true degenerecies caused by the source models: sky positions of spinless MBH are degenerate.

Although Challenge 2 saw a few successful detections of EMRI signals, data sets contained in Challenge 1B represented the first real testbed for the search algorithms developed for this critical source class.

Entries were received from three groups that used time-frequency methods and developed different coherent approaches based on Monte Carlo techniques, respectively.

Both the time-frequency and coherent approaches succeeded in detecting the relatively strong EMRI signals and in constraining their parameters, although not all groups analyzed all data sets, and the performance of the same search pipeline varied across them. The obtained results indicate that the main challenge for (isolated) EMRI analyses is the very complex structure of the likelihood surface in source-parameter space, which features a number of secondary maxima of similar height, even more so than for MBH binaries.

It would be inappropriate at this time to draw general conclusions about the relative merits of search methods and about the expected science payoff of LISA EMRI astronomy: it is not known how these techniques scale as the SNR decreases and in situations where EMRI signals overlap with each other and are affected by Galactic confusion noise. The first two complications will be addressed in Challenge 3.

8.2.4. Challenge 3

Challenge 3 builds upon Challenge 2, using improved source models and some new sources; it is due to end in April 2009. A detailed description of Challenge 3 can be found in (Babak et al., 2008b).

The data sets for Challenge 3 are released both as time series of equivalent strain generated by the LISA Simulator (The LISA Simulator) and as time series of fractional frequency fluctuations generated by Synthetic LISA (Synthetic LISA). The Challenge 3 data sets are built using the "pseudo-LISA" model of Challenges 1 and 2: the orbits of the LISA spacecraft are e^2-accurate Keplerian ellipses with conventional orientations and time offsets; *modified* TDI (a.k.a. TDI 1.5) expressions are used for the observables; and Gaussian, stationary instrument noise is included from six proof masses and six optical benches with known noise levels that are identical across each set of six.

- Galactic GW foreground from ~ 60 million compact binary systems. This data set is a direct descendant of Challenge 2, but it improves on the realism by including both detached and interacting binaries with intrinsic frequency drifts (either positive or negative).

- GW signals from 4–6 binaries of spinning MBHs, on top of a confusion Galactic-binary background. This data set improves on the realism of Challenges 1 and 2 by modeling the orbital precession (and ensuing GW modulations) due to spin-orbit and spin-spin interactions.

  Because this challenge focuses on the effects of spins rather than on the joint search for MBH signals and for the brightest Galactic binaries, the background is already *partially subtracted*—it is generated from the population of detached binaries used for Challenge 3, withholding all signals with an SNR > 5.

- GW signals from five EMRI. This challenge introduces the complication of detecting five such signals with lower SNRs, and *in the same data set*. By contrast, Galactic confusion is not included.

- GW burst signals from cosmic string cusps, occurring as a Poissonian random process throughout the data set, with a mean of five events.

- Stochastic GW background, isotropic, unpolarized, Gaussian and stationary; its spectrum grows at low frequencies as $1/f^3$, and its magnitude is set to a few times the secondary noise over a broad range of frequencies.

8.3. Source modeling

There are still open issues on the gravitational modeling side that need to be solved by the time LISA flies to ensure *maximum* science exploitation of the data. However, not only has steady progress been made on all the fronts in the past several years, in some cases the advances have been positively spectacular! Furthermore, the community behind this effort is continuously growing, to a large extent driven by the ground-based observational program that has now reached maturity. So we can be confident that the necessary waveform generation machinery will be in place for LISA.

For the overwhelming majority of the LISA sources (stellar-mass compact binaries), waveforms—based on Taylor expansions of the phase evolution—are in hand *today* and no additional theoretical work is required to enable LISA observations of this class of objects. However, a fully satisfactory physical interpretation of the LISA results will require further modeling work of the evolution of binaries under the effect of radiation reaction *and* mass transfer and/or tidal effects. Despite being a challenging problem, it has several (and more pressing) implications in a number of other fields in astronomy and astrophysics (besides gravitational-wave observations) and is actively pursued by a large theoretical astrophysics community.

There are two main open problems in LISA data analysis. The first one is the computation of waveforms produced over the very last few cycles of inspiral and during the merger of black-holes binaries (which requires a fully relativistic approach to the two body problem). This includes SMBH waveforms in PN-approximation that would include all the effects: spin-orbital coupling, eccentricity and higher order corrections to the amplitude. Those complete models have not much relevance in ground-based gravitational-wave astronomy, but might be important for LISA. This point was realized in the scientific community and several groups have started work on producing a complete waveform and assessing relevance of various effects in LISA data analysis.

The other open problem is the computation of exact waveforms for extreme mass ratio inspirals, inclduing the spin of the central object as well as eccentric orbits.

Black hole mergers

As far as black hole mergers are concerned, numerical relativity – Einstein's field equations solved directly on a supercomputer – is now advancing at a dramatic pace, with major breakthroughs in the last years: three groups working independently (Baker et al., 2006a,b, 2007b, 2008; Buonanno et al., 2007a; Campanelli et al., 2006a,b,c,d, 2008; Gonzalez et al., 2007; Pan et al., 2008; Pretorius, 2005) have demonstrated robust results for the last several cycles (< 10) of gravitational radiation generated in a variety of merger configurations. These simulations continue to improve, with a notable swath of parameter space already studied through simulations of the last few orbits.

To show consistency, the numerically obtained waveforms for non-spinning binaries have been compared to PN waveforms and show good agreement (Ajith et al., 2007; Baker et al., 2006a, 2007a; Berti et al., 2007; Buonanno et al., 2007a; Campanelli et al., 2008; Hannam et al., 2008; Pan et al., 2008) and first results with spin have been obtained (Brügmann et al., 2008; Lousto and Zlochower, 2008).

Combining the PN waveforms and the numerically obtained waveforms allows to create one waveform spanning all of the inspiral, merger and ringdown phases of the MBH (Ajith et al., 2007; Buonanno et al., 2007a,b; Pan et al., 2008), saving considerably on the computational cost for waveforms.

Even more economical are phenomenological parametrised waveforms, that combine PN waveforms and numerically obtaiend waveforms to build an *analytic* model that matches the true waveform sufficiently well (Ajith et al., 2007, 2008; Buonanno et al., 2007b).

A different formulation of the PN formalism leads to EOB waveforms (Buonanno et al., 2007b; Damour, 2001; Damour and Nagar, 2007; Damour et al., 2008a,b) that have been compared against numerical results with good agreement (Damour and Nagar, 2008).

There is indeed good reason for optimism that the full coalescence waveforms, i. e. arbitrary spins and mass ratios, will be available for LISA data analysis in the near future.

Extreme mass ratio inspirals

Considerable progress has also been made on waveform computations for EMRIs. Over timescales of order a day, the orbits of the smaller body are essentially geodesics in the spacetime of the MBH. The computational machinery for calculating the EMRI waveforms on this timescale is *already* in place (Drasco and Hughes, 2006). To correctly model the orbit over timescales longer than ~ 1 day, one must include the effect of radiation reaction, which causes the geodesic constants of integration to slowly evolve as the system radiates away energy and angular momentum. The natural approach to this problem is to take advantage of the tiny mass ratio m_2/m_1 and use perturbation theory: that is, treat the mass of the smaller body as the source of a metric perturbation on the spacetime of the MBH. A general (and quite beautiful) prescription to tackle this problem was given independently by Mino et al. (1997) and Quinn and Wald (1997), but implementing their prescription numerically has been a challenge. How to do so is at once a fascinating question in pure general relativity and a practical requirement for deriving the maximum science possible from LISA data. Recently, Mino (2005) has suggested a relatively simple and practical solution to this problem, based on calculating the average effect of radiation reaction over many orbits rather than the instantaneous self-force. Based on this solution, an explicit computational procedure for obtaining waveforms has been described (Drasco et al., 2005; Sundararajan et al., 2008). Going beyond the adiabatic apporach, a technique based on two-timescale analysis has been demonstrated, allowing to assess the higher order self force corrections (Hinderer and Flanagan, 2008).

An alternative approach, based on a mode-sum numerical procedure for regularizing the self-force has been developed (Barack and Ori, 2003), and its effectiveness has been demonstrated for circular orbits in Schwarzschild, where the results can be checked using simpler methods (Barack and Lousto, 2005). This approach can be extended to Kerr black holes (Barack et al., 2007).

Therefore it seems very likely that theorists will provide extremely accurate EMRI template waveforms well before LISA actually taking data.

Besides those very accurate waveforms based on the Teukolsky formalism (Teukolsky, 1973), several other approximate ("kludge") waveforms were suggested (Babak et al., 2007; Barack and Cutler, 2004; Gair and Glampedakis, 2006; Gair et al., 2005). Those waveforms are relatively fast to generate and they resemble the signal on the scale of few month well enough to be used in the semi-coherent search.

8.4. Heritage

The LISA data set consists of a number of time series—the TDI observables—that contain full information about the sources through the amplitude and phase evolution of waveforms, which need to be untangled through suitable data analysis. Indeed, LISA data analysis has its roots in a field with a long and venerable history: time-series analysis. There are large communities active in this field, both in theory and applications. More recently, the progress of ground-based gravitational-wave experiments has led to the specific development and implementation of data-analysis techniques for gravitational-wave observations and there is now a substantial community world-wide (several hundreds of scientists) active in this area.

As of this writing, several ground-based laser interferometers are running continuously to observe the gravitational-wave sky. The enormous increase in the amount and quality of data—the LIGO interferometers are at design sensitivity since the end of 2005 (Waldman and the LIGO Science Collaboration, 2006), and GEO and VIRGO are rapidly reaching this target (Acernese et al., 2006; Lück et al., 2006)—has led to the development and implementation of data-analysis techniques and infrastructure for gravitational-wave astronomy. Such techniques provides a solid ground on which a LISA specific data-analysis approach can be developed. Of course, the emphasis of ground-based observations is on detection of weak and rare signals contained in large data sets, but in future the parameter-estimation aspect of the analysis will become more important as sources are discovered. However, the vast majority of source-specific problems (binary inspirals, quasi-monochromatic gravitational waves, bursts and transients, stochastic backgrounds) and typical analysis issues arising from real data (non stationarity, non Gaussianity, instrument drop-out, gaps, etc.) are now being tackled routinely. A number of analysis pipelines to process, both on-line and off-line, the data sets from multiple detectors with a variety of approaches are now available (Acernese et al., 2006; Dietz et al., 2006). The analysis of LISA data sets brings in new challenges (such as overlapping sources and accurate parameter estimations) and new source classes, such as the extreme-mass ratio inspirals. However, the progress of gravitational-wave data analysis as a whole over the past ten years, and the LISA-specific effort that is now taking place give us every reasons to be optimistic about the success of the development of the necessary techniques and infrastructure for the mission.

8.5. Community Resources

During the last couple of years, a number of tools have been devloped by the LISA community to aid LISA Data Analysis in various ways. Among those tools are available to the public are three different simuators for LISA LISACode; Synthetic LISA; The LISA Simulator as well as an Online Sensitivity Curve Generator and a web based calculator for the expected SNR of different sources (The LISA Calculator).

In the framework of the MLDC, the codes and the pipelines for creating the waveforms used in the challenges have been made public (see MLDC webpage) as well as a document summarizing the conventions used in the challenges (MLDC omnibus). Further simulated data sets as well as the related software can be found on the Testbed for LISA Data Analysis.

8.6. Activities coordinated by the agencies

8.6.1. ESA-led activities

ESA issued a call for letters of intent in May 2005 to assess the interest in the European scientific community regarding LISA Data Analysis. The letter received a strong response, with responses from 8 national collaborations, representing more than 40 groups with a total of 250 researchers involved. The representatives of the national collaborations form the ESA appointed Data Analysis Steering Team (DAST), which maintains and executes the Data Analyis Plan for LISA. A preliminary plan, outlining roles and responsibilities as well as a schedule for research activities in Europe has been created and published end of 2005. Currently, the national groups do not receive funding from ESA for LISA Data Analysis; adherence to the plan is under the responsibility of the national groups.

Almost all of the planned research activities and tasks are now covered by the national groups' participation in the MLDC.

8.6.2. NASA-led activities

NASA created a plan to cover Analysis Methods for Interferometric Gravitational-wave Observations from Space (AMIGOS). The plan describes work areas, priorities, levels of effort, and timelines for the development of methods, tools, and other resources for the scientific analysis of LISA data. The structure and level of planning detail in this document parallel those of the LISA Project's Technology Development Plan. During the Project Formulation Phase, AMIGOS will be an evolving document, held by the LISA Project, and coordinated with analogous ESA plans (see section 8.6.1). It will help guide NASA-supported efforts in LISA data analysis between now and LISA's launch.

8.7. References

Xspec, an X-ray Spectral Fitting Package.

F. Acernese, P. Amico, M. Al-Shourbagy, S. Aoudia, S. Avino, et al. The status of VIRGO. *Class. Quantum Grav.*, 23(8):S63–S69, 2006.

F. Acernese, P. Amico, M. Alshourbagy, S. Aoudia, S. Avino, et al. The status of coalescing binaries search code in Virgo, and the analysis of C5 data. *Class. Quantum Grav.*, 23:187–+, April 2006.

P. Ajith, S. Babak, Y. Chen, M. Hewitson, B. Krishnan, et al. A phenomenological template family for black-hole coalescence waveforms. *Classical and Quantum Gravity*, 24:689–+, October 2007, also as arXiv:0704.3764.

P. Ajith, S. Babak, Y. Chen, M. Hewitson, B. Krishnan, et al. Template bank for gravitational waveforms from coalescing binary black holes: Nonspinning binaries. *Phys. Rev. D*, 77(10):104017–+, May 2008, also as arXiv:0710.2335.

K. A. Arnaud, G. Auger, S. Babak, J. G. Baker, M. J. Benacquista, et al. Report on the first round of the Mock LISA Data Challenges. *Class. Quantum Grav.*, 24:529–+, October 2007a, also as arXiv:gr-qc/0701139.

K. A. Arnaud, S. Babak, J. G. Baker, M. J. Benacquista, N. J. Cornish, et al. An overview of the second round of the Mock LISA Data Challenges. *Class. Quantum Grav.*, 24:551–+, October 2007b, also as arXiv:gr-qc/0701170.

K. A. Arnaud, S. Babak, J. G. Baker, M. J. Benacquista, N. J. Cornish, et al. (Mock LISA Data Challenge Task Force). A How-To for the Mock LISA Data Challenges. In S. M. Merkowitz and J. C. Livas, editors, *LASER INTERFEROMETER SPACE ANTENNA: 6th International LISA Symposium*, volume 873, pages 625–632. AIP, 2006, also as arXiv:gr-qc/0609106.

S. Babak, H. Fang, J. R. Gair, K. Glampedakis, and S. A. Hughes. "Kludge" gravitational waveforms for a test-body orbiting a Kerr black hole. *Phys. Rev. D*, 75(2):024005–+, January 2007, also as arXiv:gr-qc/0607007.

S. Babak, J. G. Baker, M. J. Benacquista, N. J. Cornish, J. Crowder, et al. Report on the second Mock LISA data challenge. *Class. Quantum Grav.*, 25(11):114037–+, June 2008a, also as arXiv:0711.2667.

S. Babak, J. G. Baker, M. J. Benacquista, N. J. Cornish, J. Crowder, et al. The Mock LISA Data Challenges: from Challenge 1B to Challenge 3. *Class. Quantum Grav.*, 25(18):184026–+, September 2008b, also as arXiv:0806.2110.

J. G. Baker, J. Centrella, D.-I. Choi, M. Koppitz, and J. van Meter. Binary black hole merger dynamics and waveforms. *Phys. Rev. D*, 73(10):104002–+, May 2006a, also as arXiv:gr-qc/0602026.

J. G. Baker, J. Centrella, D.-I. Choi, M. Koppitz, and J. van Meter. Gravitational-Wave Extraction from an Inspiraling Configuration of Merging Black Holes. *Physical Review Letters*, 96(11):111102–+, March 2006b, also as arXiv:gr-qc/0511103.

J. G. Baker, J. R. van Meter, S. T. McWilliams, J. Centrella, and B. J. Kelly. Consistency of Post-Newtonian Waveforms with Numerical Relativity. *Physical Review Letters*, 99(18):181101–+, November 2007a, also as arXiv:gr-qc/0612024.

J. G. Baker, J. R. van Meter, S. T. McWilliams, J. Centrella, and B. J. Kelly. Consistency of post-Newtonian waveforms with numerical relativity. *Physical Review Letters*, 99(18):181101–+, November 2007b, also as arXiv:gr-qc/0612024.

J. G. Baker, W. D. Boggs, J. Centrella, B. J. Kelly, S. T. McWilliams, et al. Mergers of nonspinning black-hole binaries: Gravitational radiation characteristics. *Phys. Rev. D*, 78(4):044046–+, August 2008, also as arXiv:0805.1428.

L. Barack and C. Cutler. LISA capture sources: Approximate waveforms, signal-to-noise ratios, and parameter estimation accuracy. *Phys. Rev. D*, 69:082005, April 15 2004, also as arXiv:gr-qc/0310125.

L. Barack and C. O. Lousto. Perturbations of Schwarzschild black holes in the Lorenz gauge: Formulation and numerical implementation. *Phys. Rev. D*, 72(10): 104026–+, November 2005.

L. Barack and A. Ori. Gravitational Self-Force on a Particle Orbiting a Kerr Black Hole. *Physical Review Letters*, 90(11):111101, March 2003, also as arXiv:gr-qc/0212103.

L. Barack, D. A. Golbourn, and N. Sago. m-mode regularization scheme for the self-force in Kerr spacetime. *Phys. Rev. D*, 76(12):124036–+, December 2007, also as arXiv:0709.4588.

E. Berti, V. Cardoso, J. A. Gonzalez, U. Sperhake, M. Hannam, et al. Inspiral, merger, and ringdown of unequal mass black hole binaries: A multipolar analysis. *Phys. Rev. D*, 76(6):064034–+, September 2007, also as arXiv:gr-qc/0703053.

B. Brügmann, J. A. González, M. Hannam, S. Husa, and U. Sperhake. Exploring black hole superkicks. *Phys. Rev. D*, 77(12):124047–+, June 2008, also as arXiv:0707.0135.

A. Buonanno, G. B. Cook, and F. Pretorius. Inspiral, merger and ring-down of equal-mass black-hole binaries. *Phys. Rev. D*, 75(12):124018–+, June 2007a, also as arXiv:gr-qc/0610122.

A. Buonanno, Y. Pan, J. G. Baker, J. Centrella, B. J. Kelly, et al. Approaching faithful templates for nonspinning binary black holes using the effective-one-body approach. *Phys. Rev. D*, 76(10):104049–+, November 2007b, also as arXiv:0706.3732.

M. Campanelli, C. O. Lousto, P. Marronetti, and Y. Zlochower. Accurate Evolutions of Orbiting Black-Hole Binaries without Excision. *Physical Review Letters*, 96(11): 111101–+, March 2006a, also as arXiv:gr-qc/0511048.

M. Campanelli, C. O. Lousto, and Y. Zlochower. Last orbit of binary black holes. *Phys. Rev. D*, 73(6):061501–+, March 2006b, also as arXiv:gr-qc/0601091.

M. Campanelli, C. O. Lousto, and Y. Zlochower. Spinning-black-hole binaries: The orbital hang-up. *Phys. Rev. D*, 74(4):041501–+, August 2006c, also as arXiv:gr-qc/0604012.

M. Campanelli, C. O. Lousto, and Y. Zlochower. Spin-orbit interactions in black-hole binaries. *Phys. Rev. D*, 74(8):084023–+, October 2006d, also as arXiv:astro-ph/0608275.

M. Campanelli, C. O. Lousto, H. Nakano, and Y. Zlochower. Comparison of Numerical and Post-Newtonian Waveforms for Generic Precessing Black-Hole Binaries. *ArXiv e-prints*, August 2008, also as arXiv:0808.0713.

T. Damour. Coalescence of two spinning black holes: An effective one-body approach. *Phys. Rev. D*, 64(12):124013–+, December 2001, also as arXiv:gr-qc/0103018.

T. Damour and A. Nagar. Faithful effective-one-body waveforms of small-mass-ratio coalescing black hole binaries. *Phys. Rev. D*, 76(6):064028–+, September 2007, also as arXiv:0705.2519.

T. Damour and A. Nagar. Comparing effective-one-body gravitational waveforms to accurate numerical data. *Phys. Rev. D*, 77(2):024043–+, January 2008, also as arXiv:0711.2628.

T. Damour, A. Nagar, E. N. Dorband, D. Pollney, and L. Rezzolla. Faithful effective-one-body waveforms of equal-mass coalescing black-hole binaries. *Phys. Rev. D*, 77(8):084017–+, April 2008a, also as arXiv:0712.3003.

T. Damour, A. Nagar, M. Hannam, S. Husa, and B. Brügmann. Accurate effective-one-body waveforms of inspiralling and coalescing black-hole binaries. *Phys. Rev. D*, 78(4):044039–+, August 2008b, also as arXiv:0803.3162.

A. Dietz et al. (LIGO Scientific Collaboration). Overview of LIGO Scientific Collaboration inspiral searches. *Class. Quantum Grav.*, 23:705–+, October 2006.

S. Drasco and S. A. Hughes. Gravitational wave snapshots of generic extreme mass ratio inspirals. *Phys. Rev. D*, 73(2):024027, January 2006.

S. Drasco, É. É. Flanagan, and S. A. Hughes. Computing inspirals in Kerr in the adiabatic regime: I. The scalar case. *Class. Quantum Grav.*, 22:801–+, August 2005, also as arXiv:gr-qc/0505075.

J. R. Gair and K. Glampedakis. Improved approximate inspirals of test bodies into Kerr black holes. *Phys. Rev. D*, 73(6):064037–+, March 2006, also as arXiv:gr-qc/0510129.

J. R. Gair, D. J. Kennefick, and S. L. Larson. Semirelativistic approximation to gravitational radiation from encounters with nonspinning black holes. *Phys. Rev. D*, 72 (8):084009–+, October 2005, also as arXiv:gr-qc/0508049.

J. A. Gonzalez, U. Sperhake, B. Bruegmann, M. Hannam, and S. Husa. Total recoil: the maximum kick from nonspinning black-hole binary inspiral. *Physical Review Letters*, 98(9):091101–+, March 2007, also as arXiv:gr-qc/0610154.

M. Hannam, S. Husa, J. A. González, U. Sperhake, and B. Brügmann. Where post-Newtonian and numerical-relativity waveforms meet. *Phys. Rev. D*, 77(4): 044020–+, February 2008, also as arXiv:0706.1305.

I. W. Harry, S. Fairhurst, and B. S. Sathyaprakash. A hierarchical search for gravitational waves from supermassive black hole binary mergers. *Class. Quantum Grav.*, 25(18):184027–+, September 2008, also as arXiv:0804.3274.

T. Hinderer and É. É. Flanagan. Two-timescale analysis of extreme mass ratio inspirals in Kerr spacetime: Orbital motion. *Phys. Rev. D*, 78(6):064028–+, September 2008, also as arXiv:0805.3337.

LISA International Science Team. Science requirements document, December 2006.

LISACode. LISACode. online at `http://www.apc.univ-paris7.fr/APC_CS/Animation/Reseaux/LISA-France/analyse.phtml`.

C. O. Lousto and Y. Zlochower. Further insight into gravitational recoil. *Phys. Rev. D*, 77(4):044028–+, February 2008, also as arXiv:0708.4048.

H. Lück, M. Hewitson, P. Ajith, B. Allen, P. Aufmuth, et al. Status of the GEO600 detector. *Class. Quantum Grav.*, 23(8):S71–S78, 2006.

Y. Mino. Extreme mass ratio binary: radiation reaction and gravitational waveform. *Class. Quantum Grav.*, 22:375–+, May 2005, also as arXiv:gr-qc/0506008.

Y. Mino, M. Sasaki, and T. Tanaka. Gravitational radiation reaction to a particle motion. *Phys. Rev. D*, 55:3457–3476, March 1997, also as arXiv:gr-qc/9606018.

MLDC omnibus. Document for Challenge 1, 2006. online at `http://svn.sourceforge.net/viewvc/lisatools/Docs/challenge1.pdf`.

MLDC webpage. Mock LISA Data Challenges, 2006. online at `http://astrogravs.nasa.gov/docs/mldc/`.

MLDC Wiki. MLDC Task Force Wiki, 2006. online at `http://www.tapir.caltech.edu/dokuwiki/listwg1b:home`.

Online Sensitivity Curve Generator. Online Sensitivity Curve Generator. online at `http://www.srl.caltech.edu/~shane/sensitivity/`.

Y. Pan, A. Buonanno, J. G. Baker, J. Centrella, B. J. Kelly, et al. Data-analysis driven comparison of analytic and numerical coalescing binary waveforms: Nonspinning case. *Phys. Rev. D*, 77(2):024014–+, January 2008, also as arXiv:0704.1964.

F. Pretorius. Evolution of Binary Black-Hole Spacetimes. *Physical Review Letters*, 95 (12):121101–+, September 2005, also as arXiv:gr-qc/0507014.

T. C. Quinn and R. M. Wald. Axiomatic approach to electromagnetic and gravitational radiation reaction of particles in curved spacetime. *Phys. Rev. D*, 56: 3381–3394, September 1997, also as arXiv:gr-qc/9610053.

P. A. Sundararajan, G. Khanna, S. A. Hughes, and S. Drasco. Towards adiabatic waveforms for inspiral into Kerr black holes. II. Dynamical sources and generic orbits. *Phys. Rev. D*, 78(2):024022–+, July 2008, also as arXiv:0803.0317.

Synthetic LISA. Synthetic LISA. online at http://www.vallis.org/syntheticlisa/.

Testbed for LISA Data Analysis. Testbed for LISA Data Analysis. online at http://tla.gravity.psu.edu/index.shtml.

S. A. Teukolsky. Perturbations of a Rotating Black Hole. I. Fundamental Equations for Gravitational, Electromagnetic, and Neutrino-Field Perturbations. *ApJ*, 185: 635–648, October 1973.

The LISA Calculator. The LISA Calculator. online at http://www.physics.montana.edu/LISA/lisacalculator/.

The LISA Simulator. The LISA Simulator. online at http://www.physics.montana.edu/lisa/.

S. J. Waldman and the LIGO Science Collaboration. Status of LIGO at the start of the fifth science run. *Class. Quantum Grav.*, 23:3–+, October 2006.

9. Outline of the Data-Analysis Architecture

At the present stage of the LISA project (Formulation), the Data-Analysis Architecture has not yet been finalized, as some details of the architecture will depend on the outcome of ongoing studies, such as the problem of global fitting.

Therefore, this chapter describes the general outline of the Data-Analysis Architecture as it is foreseen at this point in time.

9.1. Observation modes

As LISA is an omnidirectional instrument, it requires no pointing or scheduling of observation time. With the exception of scheduled interruptions (*e. g.*, for antenna re-pointing) LISA will observe all the sky all the time.

To allow for *uninterrupted* observations of particular events, such as the last phase of a SMBH merger, it will be possible to disallow scheduled interruptions in a "protected period".

9.2. Data rates and downlink

The net effective continuous data rate for the constellation[1] is 15 kbps, of which 5 kbps is housekeeping data and 10 kbps is science data.

Nominal communication with the LISA constellation will be conducted through the DSN 34 m network with the 70 m network as a backup. The communication schedule foresees 8 hours of contact every other day with a actual bit rate of 90 kbps, resulting in an average *daily* data volume of 1.3 GByte and a total data volume for the nominal observation time (5 years) of the mission of about 2.4 TByte of which about 800 GByte are science data (Diehl, 2009).

With these (compared to other astronomy missions quite small) data rates, no on-board reduction of the data will take place, *i. e.* all the data taken on the spacecraft will be available on ground, not just the science data.

Each satellite will be contacted every 6 days, which implies inter-spacecraft communication to relay the data obtained at each spacecraft to the communicating spacecraft. The backup scenario foresees direct communication which each spacecraft with on-board storage of the obtained data.

[1] This is the rate which which the constellation is amassing data. It is not to be confused with the *downlink* rate

In cases of impending merger events, additional data downloads can be scheduled to get updates on source parameters (*e. g.* sky position) and facilitate coordination with electro-magnetic observations.

9.3. Data format

Science Data The interferometry on LISA comprises three interferometers on each optical bench (of which there are two per spacecraft). The main interferometer (*i. e.*, the one that combines the incoming light from the remote satellite with a local oscillator) and the proof mass interferometer (*i. e.*, the one that is used to asses the position of the proof mass of the inertial sensor with respect to the optical bench) are read out via a quadrant photodiode to assess the relative angle between the interfering light beams. The reference interferometer, used to assess the local oscillator is read out via a single-element photodiode, as this interferometer does not contain any moving parts.

The ancillary measurements comprise the measurement of the clock noise and the armlength, both achieved through a sideband modulation–demodulation scheme.

Thus, the LISA science data consists of the following

- 6 main interferometric measurements, 4 channels each
- 6 proof mass measurements, 4 channels each
- 6 reference interferometer, 1 channel each
- 6 measurements of the ancillary measurements, 1 channel each

The bandwidth for the science data is 1 Hz, the sampling rate is 3 Hz, the bit depth varies depending on the required dynamical range between 10 bit and 24 bit.

Science housekeeping data These consist of continuously measured environmental data (temperatures, magnetic field, proton radiation) and of the data from the sensors used in the Drag Free and Attitude Control System (DFACS). Both types of data are low-pass–filtered and represented as a continuously sampled data stream at low bandwidth (typically lower than 0.1 Hz). Science housekeeping data can be used to improve the main science data product by removing systematic errors and by flagging time periods during which unusual environmental conditions where present. It is anticipated that the statistical evaluation of archived science housekeeping data collected over a longer period will be necessary to test and calibrate the algorithms intended to improve the basic science data. Therefore these algorithms will not be part of baseline data processing.

Technical Housekeeping data These consist of status data from the technical equipment of spacecraft and payload, as needed to identify malfunctions or parametric degradation. For the purpose of analyzing complex functions, such as internal control loops, bursts of high-speed, high-resolution sensor data collected over a short time period may be included. Other than for science housekeeping, acquired sensor data in the technical-housekeeping data stream are not intended to form a continuously recorded data stream represented with proper sampling and anti-aliasing filtering. A high flexibility in selecting housekeeping data objects and their sampling rate by ground command is foreseen to support monitoring and diagnosis of the payload, while maintaining moderate average data rates. On-board recording of housekeeping data (at increased average data rates) in a circular buffer capable of storing several weeks' worth of data is foreseen to further enhance the usefulness of technical housekeeping for investigation of anomalies.

Once on the ground, the science data will be processed (see Appendix A) to produce data streams that represent the strain $h(t)$ as measured by LISA.

9.4. Data processing

LISA Data Analysis requires a significant amount of processing power. Both the production of source templates as well as the extraction of the data from the data are numerically intensive. Currently, supercomputers are used for the calculation of SMBH waveforms, *e. g.* through Project Columbia whereas other waveforms can be calculated in a few hours on a moderately fast workstation.

Given the steady increase in computing power in the past and assuming that this increase will go on for some time, LISA Data Analysis will not put exceptional requirements on future computational facilities.

9.5. Data Products

The guiding principle for publication of the data products is to enable the scientific community to re-do any analysis of the data, starting from the Level 0 products to Level 3 products. It is therefore required to not limit the publication to the data, but to also make available the algorithms, the software, and the models used for processing the data as well as ansuring that the data processing history for any data published by the LSDC is traceable and retrievable.

Thus, the LSDC will make available the following science products to the scientific community at large.

Source catalog (Level 3) One of the main products of the LISA Science Data Centre (LSDC) will be the publication of a source catalog, containing the identified sources, their physical and astrophysical parameters (including confidence intervals or probability density functions), potential electro-magnetic counterparts, as well as their strain time series $h(t)$. The source catalog will be open to the public and accessible

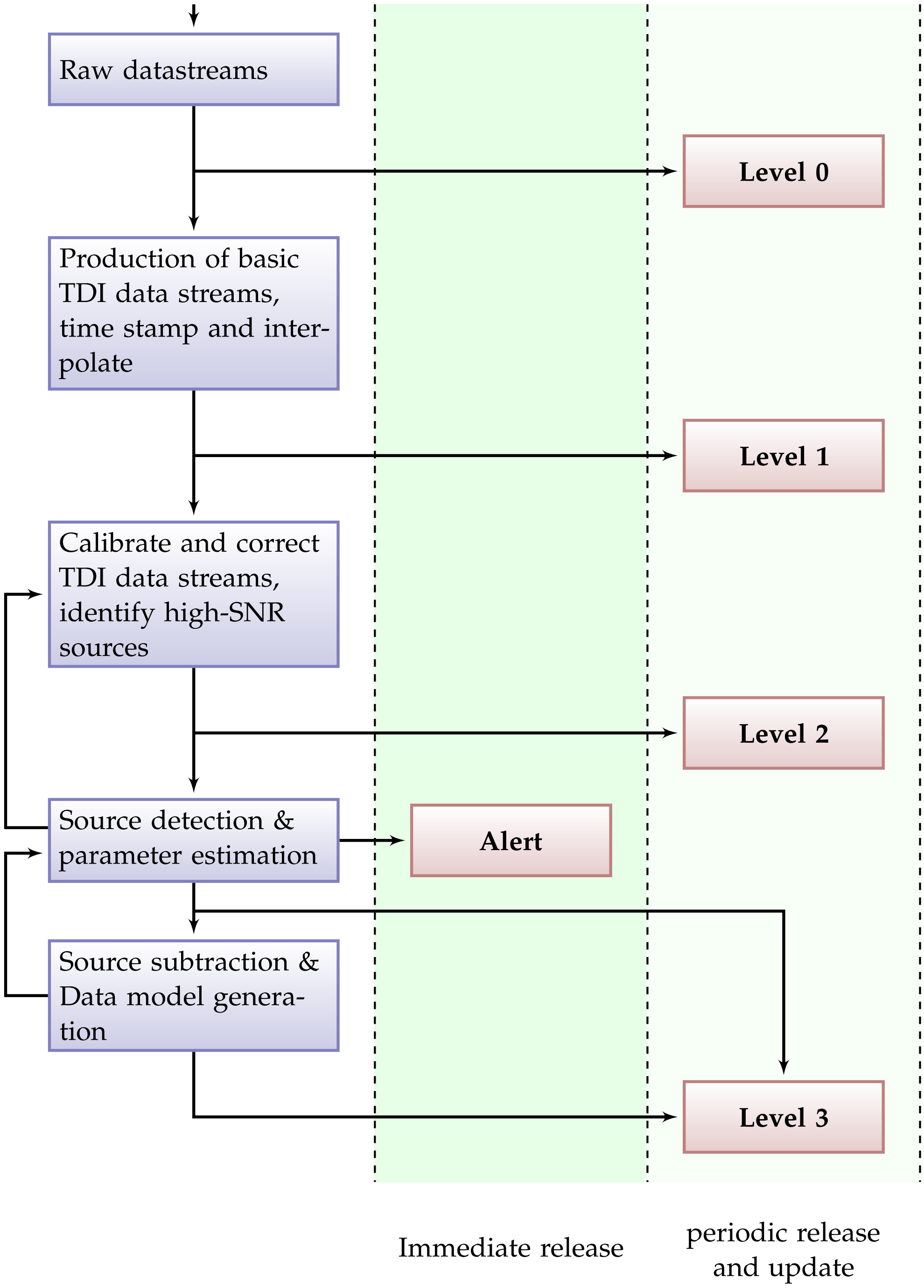


Figure 9.1.: LISA data flow. The leftmost column shows in blue the processing steps from reception of the raw data from MOC to the final source detection, the data flow is indicated with arrows. The different data products are given in red. The alerts are released immediately after they are known, whereas the other data products are released according to a release and update schedule.

through common interfaces; it will be updated regularly until it includes all the results obtained by the mission.

Processed data streams (Level 2) Level 2 data comprises the fully processed data streams that are needed to isolate individual gravitational wave signals using parametrized source models or other data analysis techniques. Level 2 products are fully public.

The Level 2 data consists of fully calibrated and corrected TDI data streams, augmented by the spacecraft ephemerides and data streams that contain the current best estimates for cataloged signals (Level 3) that can be used, *e. g.* to subtract the known signals prior to further data analysis.

In addition, Level 2 data contains the software and the models used to produce Level 3 data from Level 2 data, in particular the software for modeling the instrument response to gravitational waves and the current best estimate for the noise spectra of the instrument.

Basic data streams (Level 1) Level 1 data comprises the data streams necessary to obtain the fully calibrated and corrected TDI data streams. Level 1 products are fully public.

The Level 1 data consists of the basic TDI data streams plus the relevant data streams of the GRS, the IMS, and the science housekeeping that are needed to correct the basic data streams for spurious accelerations, environmental and geometric effects.

Those data stream might include *e. g.* temperature, magnetic field, orientation of the proof masses, accumulated charge, and any information on the effective optical pathlength.

In addition, Level 1 data contains the software and the models used to produce Level 2 data from Level 1 data.

Raw data streams (Level 0) Level 0 data comprises the raw data streams necessary to obtain the basic TDI data streams and the relevant data streams of the GRS, the IMS and the science housekeeping. Level 0 products are fully public.

The Level 0 data comprise all the data from each of the phase-meter channels, all the data from the GRS, and the complete science and payload housekeeping data.

In addition, Level 0 data contains the software and the models used to produce Level 1 data from Level 0 data, in particular the full dynamical model of the GRS as well as the phase-meter algorithms.

Additional products Although not the main scientific product of the LISA mission, the measured performance and the physics model of the inertial sensor can be useful for future space missions. Therefore, the respective data will be made available in a TBD form.

A potential proprietary period, the exact scope of data rights, and the organization and funding of the LISA Science Data Centre (LSDC) will be the subject of a later MoU between ESA and NASA.

9.6. References

R. Diehl. LISA Operations Concept Document. Internal report LISA-OPS-RP-0001, LISA Project, March 2009.

10. LISA Technology Package Data Analysis

10.1. Introduction

Data analysis for LISA implictly assumes "clean" data streams, i. e. all artefacts, such as glitches, gaps or periods of decreased sensitivity, identified, removed or at least flagged, the best possible strain sensitivity achieved and quite generally having removed all traces of any underlying instrumental characteristics.

Despit the conceptional simplicity of LISA, the actual technical implementation of the mission is rather complex, resulting in a much larger number of data streams than the three TDI streams assumed for the data analysis.

Fortunately, LISA Pathfinder has to meet a similar challenge: Gather as much information from the data streams as possible and adjust the dynamical models of the payload to highest possible fidelity: Knowing the behaviour of the measurement system allows to minimize the noise in the LISA data streams.

It is therefore expected that the data analysis for LISA Pathfinder will provide LISA with wealth of knowledge and experience for the commissioning, diagnosis and operation of LISA.

The data analysis for the LISA Pathfinder mission aims to extract the maximum amount of scientific worth from the data taken during the mission. The data analysis effort takes place on two main fronts: the analysis and design of the experiments to be performed; and the development and testing of data analysis tools to be used throughout the mission life-time.

The design of the various experiments (or runs) to be performed during the mission is a critical activity which is currently underway. Each run of the mission is designed to last one or more days, with each run consisting of one or more experiments. The design of the experiments at a technical level is the task of the various scientists involved in the mission. These experiment designs (technical notes) are then converted into operating procedures which will be executed as part of the mission time-line. The collection of runs is brought together in the EMP.

The remainder of this chapter focusses on the design and development of the data analysis tools that will be used. In addition, the interaction of the tool development and the experiment design is highlighted.

10.2. Requirements and Design

The fact that the LPF mission consists of a series of around 200 discrete experiments, each potentially depending on the results of the previous experiments, places some constraints on the software used to carry out the analysis. In particular, the software must be thoroughly tested in order to avoid incorrect or erroneous decisions being taken based on incorrect analysis results. While it is envisaged that much of the analysis will be designed in advance of the mission, it is also likely that the analyses that have been developed will have to change to match any changes in the experiments performed, and to include new information and insights, as they are gained.

A further constraint on the data analysis software was put in place to ensure that the results of the data analysis are long lived, such that no information is lost between LPF and LISA. The formulation of this constraint was to require that a full processing history of any data analysis result should be kept with the result at all times.

These constraints led to the concept of the AO described in the next section.

10.2.1. Analysis Objects

The idea that an analysis result should not stand-alone and should be more that a numeric data series, or a plot, led to the concept of an Analysis Object, or AO. An AO contains a full processing history of everything that has happened to it, from the moment it was created (from raw data, for example), to the point where the analysis concludes and the result is stored. In addition, various other pieces of meta-information are kept in the AO, for example, details of who created the object. Figure 10.1 shows a schematic of an AO.

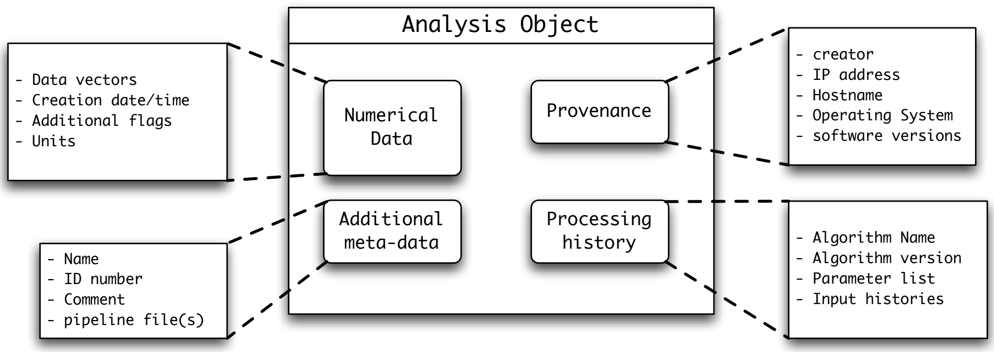


Figure 10.1.: Schematic of the structure of an Analysis Object. The diagram expands the first layer of an AO to reveal some details of the underlying structure.

In order to ensure this 'history tracking', the algorithms used to process the data must be 'history aware'. In other words, they must extend the history tree of the inputs and attach this extended history to their outputs. A schematic of this process is show in Figure 10.2.

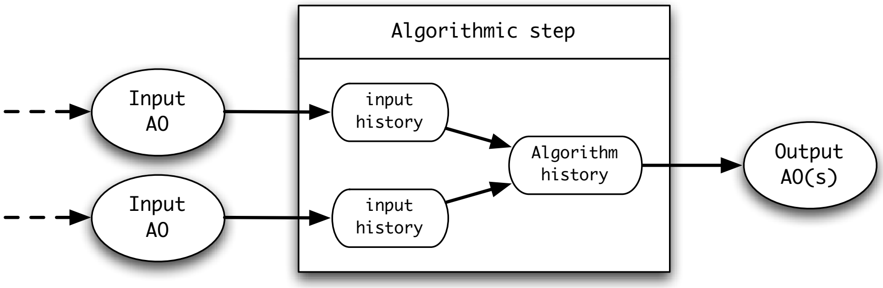


Figure 10.2.: Schematic of the history capturing concept for algorithms working on Analysis Objects.

On application of each algorithm in an analysis pipeline, the algorithm adds a *history step* which contains:

- the name of the algorithm,
- the version of the algorithm,
- the configuration parameters for the algorithm,
- the names of the input objects,
- the histories associated with each input object.

In this way, a processing tree is built-up as an object passes through an analysis pipeline.

Once all this information is collected it can be extracted from the object and used to, for example, view the processing history, rebuild the object following the same processing steps, or reproduce the processing pipeline for alteration. An example of the processing tree of a particular analysis, automatically generated from the resulting object, is shown in Figure 10.3.

10.2.2. Implementation

The current implementation of this conceived analysis environment is done in MATLAB. This choice was made in order to reduce the amount of testing needed on the software, such that existing MATLAB routines don't undergo the full rigorous testing that is applied to purpose written algorithms.

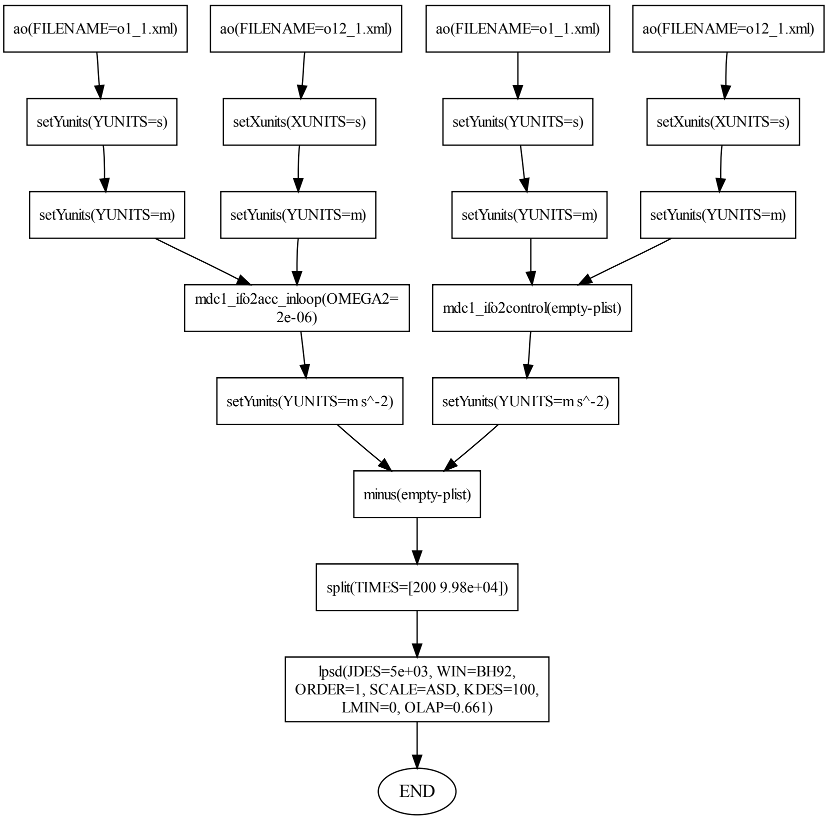


Figure 10.3.: History tree of MDC1 signal processing pipeline. The tree is produced automatically by calling the command `dotview` on the AO history object.

Using an object oriented approach, the AO concept is implemented based on a set of user classes and supporting classes. The main user classes are described in Table 10.1.

In this infrastructure, all processing algorithms are methods of one of the user classes. For example, the algorithm for estimating the spectral density of a time-series, `psd`, is a method of the AO class.

All the classes, methods, and supporting codes are brought together as a MATLAB Toolbox called LTPDA. This is freely available from http://www.lisa.aei-hannover.de/ltpda/.

10.2.3. Data storage and distribution

During the mission, telemetry data will be downloaded from the space-craft daily and must be made available to the analysis team. Typically, the raw data will be

Table 10.1.: Examples of the main user-classes implemented in the LTPDA toolbox.

Class name	description
ao	Implementation of Analysis Objects.
pzmodel	Implements S-domain pole/zero models.
parfrac	Implements partial fraction representation of transfer functions.
rational	Implements a rational representation of transfer functions.
miir	Implements IIR digital filters.
mfir	Implements FIR digital filters.
ssm	Implements state-space model objects.
plist	A class to store parameter lists. A parameter has a *key name* and a *value*.

converted in to about 2400 AOs representing the various recorded data channels acquired over the preceding 24 hours. The analysis team will then analyse the experiment(s) and produce a series of results. Over the full mission life-time, a huge number of AOs will be generated. Since each experiment will possibly take input from all previous experiments, it is necessary to have a practical way of storing and retrieving this large number of AOs. In addition, it is likely that more than one analysis will be underway at any one time. Hence, the data must be accessible concurrently from multiple analysis stations.

Due to these requirements, the LTP DA team has developed a client-server system which has a MySQL database at its core. This server part of this system is referred to as an 'LTPDA Repository'. Each repository can host a number of databases.

The LTPDA objects (AOs, *e.t.c.*), are stored directly in the database in an XML format. In addition, when an object is submitted to the database from the MATLAB client interface, various pieces of meta-data are extracted from the object and stored in different tables of the database. This provides the ability to perform complex queries on the database to recover particular results.

A web-browser-based interface has also been created for the LTPDA Repository. This serves two purposes: for administrators, this provides an easy way to manage user access permissions to the repository and its databases. For the user, this provides a convenient way to update their profile, as well as to browse the contents of the different databases in a repository.

An overview of how this system will fit into the operations center for the mission is shown in Figure 10.4.

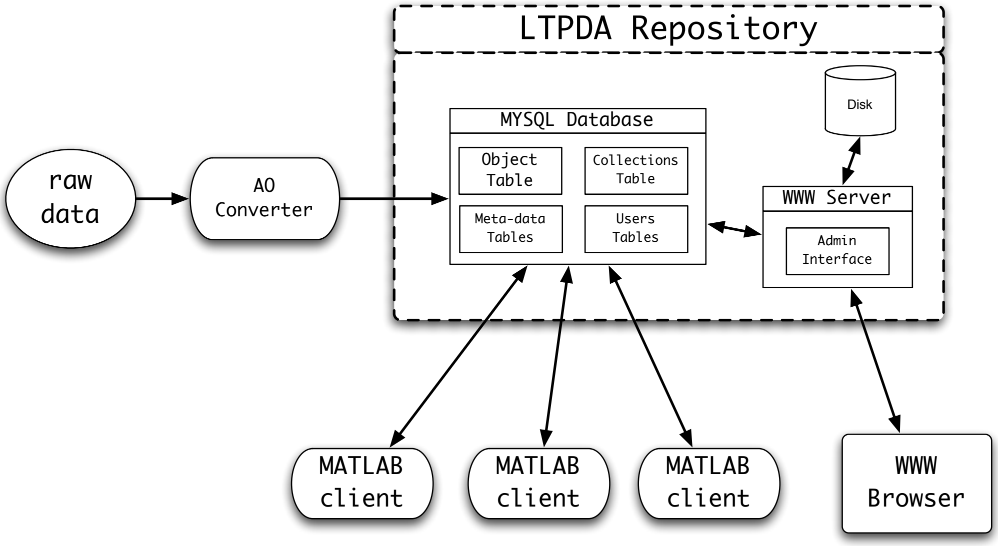


Figure 10.4.: Schematic of an LTPDA Repository set up. The core MySQL database holds all the submitted LTPDA objects, as well as detailed meta-data about each object. Interaction with the repository can be done directly from with the LTPDA toolbox, or with a standard web browser.

The use of standard, open-source software means that it is relatively easy to setup an LTPDA Repository in any location. Details of the installation process can be made available on request.

10.2.4. Mock Data Challenges

In order to help steer the design of the various algorithms for the mission, the LTP Data Analysis team is undertaking a series of MDC. Each MDC is carried out by two teams and follows a similar plan: we define a parametric model of the LTP system, which is typically simplified for the investigation we want to undertake. This model is then used by the data generation team to generate some data using particular parameter values. The details of the model, and perhaps some of the parameter values, are communicated to the data analysis team whose task it is to analyse the data in order to extract the desired results.

For example, in the first MDC, a simple 1D model of LTP was built and used to generate two time-series: one representing the displacement of the first test-mass with respect to the space-craft, the second representing the differential displacement of the two test-masses. In this case, all parameters and details of the model were communicated to the analysis team. The aim of the MDC was to take the simulated test-mass position measurements and convert them back to an equivalent out-of-loop acceleration noise. The second MDC is a continuation of MDC 1, but this time not all parameters of the model are given to the analysis team. Instead they are given a set of simulated outputs which represent different experiments. Each experiment is then to be used to estimate the parameters of the system in order to recover the best estimate of the out-of-loop differential forces on the two test-masses.

Further MDCs will build on the first two with increasing levels of complexity. The final aim is that the MDCs will approach the experiments described in the Experiment Master Plan. By doing this, we ensure that, as far as possible, all required analysis algorithms are ready for carrying out the mission.

10.3. Current Status

The development of the data analysis software is in an advanced stage. So far, one official delivery of the software (version 1.0) has been made to ESA, with another (version 2.0) expected at the beginning of 2009.

While most of the infrastructure of the data analysis environment is in place, including the full distributed data system (LTPDA Repositories), there remains an amount of work in the development of the algorithms that go into the toolbox. The toolbox already contains most of the standard signal-processing tools needed to analyse most lab-type experiments. However, as the development of the experiments continues, the development of more specific algorithms must take place. This activity goes hand-in-hand with the design and execution of the Mock Data Challenges discussed above.

10.4. Future development

By the time of the mission, the LTPDA environment, including its LTPDA Repository, will be a robust and well tested analysis environment suitable for the analysis of lab-based experiments, as well as for the analysis and characterisation of data coming from LTP and LISA.

Over the coming years, it is expected that more and more algorithms will be introduced to the toolbox, as well as additional user objects in order to represent different quantities for analysis. In particular, the experience gained from analysing the LTP data should be carried forward into LISA, particularly in terms of the results, and the tools. The suitability of a MATLAB based analysis tool for LISA is still to be assessed, though the data volumes in LISA are not significantly larger than those in LPF, so it is likely that the tools can be directly applied. However, it is also possible to port the concepts developed here to another environment, for example, Octave.

A. Interferometry and the LISA Science Analysis Pipeline

A.1. Time Delay Interferometry (TDI)

The LISA science streams will be the end product of a pipeline that will take the raw LISA data and produce a data product on which science analysis (or more colloquially, "data analysis", the subject of this document) will be carried out. This pipeline is completely analogous to the data processing in traditional electromagnetic astronomy, where raw instrumental data is calibrated and reduced to make data sets suitable for science analysis.

In LISA's case, an essential part of this pre-processing is the creation of the interferometric variables (*time-delay interferometry*, or TDI, variables) from the "raw" phase-measurement data collected aboard each LISA spacecraft.

Unlike ground-based interferometric observatories, LISA will not be a static instrument. Each individual spacecraft is on its own independent orbit, constantly in motion relative to the other spacecraft in the constellation, with arm lengths varying by as much as 1 % (Povoleri and Kemble, 2006). TDI is the cornerstone of the techniques dealing with these effects, correcting for the ever-changing armlengths of LISA (Armstrong et al., 1999), as well as for spurious noise such as optical bench motion (Estabrook et al., 2000), and clock jitter in the local oscillator frequency (Hellings, 2001a,b). A simple interpretation of the TDI procedure is that the measured phase information in each arm is delayed by a specific time, and linear combinations of the resulting data are produced in just the right way as to make the optical path lengths in the two phase signals equal. In particular, TDI combinations allow to synthesize equal-arm interferometric signals (Cornish and Hellings, 2003; Shaddock et al., 2003; Vallisneri, 2005), rendering the measurement insensitive to frequency noise of the lasers.

Detailed analyses of the measurement requirements for TDI applications have in return yielded measurement requirements of the system (Tinto et al., 2003), specifications for the ranging between spacecraft (Tinto et al., 2005), and laser stabilization (Sylvestre, 2004). Detailed modeling of TDI, using simulated data on laser telemetry, including sources of noise and realistic orbits has been demonstrated by existing LISA simulation software, notably `Synthetic LISA` (Vallisneri, 2005) and the `LISA Simulator` (Rubbo et al., 2003). Hardware realizations of the LISA TDI have been implemented and are currently being tested on the laboratory benchtop (Cruz et al., 2006; Thorpe et al., 2004; Thorpe et al., 2005). An alternative interpretation of TDI is given by Romano and Woan (2006), who connect TDI to a principal-component analysis.

Each possible TDI variable has a unique signature in any individual time-series derived from the LISA data, and is generally characterized by a *response function* $\mathcal{R}$ (also sometimes called a *transfer function*). Such response functions have been extensively discussed and analyzed in the literature (Cornish and Larson, 2001; Cornish and Rubbo, 2003; Larson et al., 2000, 2002; Rubbo et al., 2004). This response-function signature is completely analogous to the point-spread function in electromagnetic telescopes. The art of science analysis then is how to extract science from the various TDI data streams; in analogy to the procedure in the electromagnetic case, science analysis proceeds by deconvolving the response function from the data stream to yield the incident gravitational-wave field that originally impinged on the LISA observatory.

TDI data streams come in related triples, where each individual combination (heuristically speaking) corresponds to choosing one of the three LISA sciencecraft as the central element in the laser interferometer. For any given possible interferometric combination, there are two related variables that are obtained by cyclically permutating the three LISA spacecraft. In addition to the fundamental interferometric combinations (called $\{X, Y, Z\}$, similar to the familiar Michelson interferometric combination from basic optics), there are other important TDI triples that appear in the literature, including the *Sagnac combinations* $\{\alpha, \beta, \gamma\}$ (Armstrong et al., 1999), and the *optimal combinations* $\{A, E, T\}$ (Prince et al., 2002).

Since its inception, TDI has evolved through different *generations*, corresponding to the inclusion of higher-order motion effects in the interferometric time-delay corrections (Cornish and Hellings, 2003; Shaddock, 2004; Shaddock et al., 2003; Tinto et al., 2004). General results derived from first-generation TDI (Armstrong et al., 1999; Estabrook et al., 2000) carry forward and still hold true in later generations, as the corrections are small and easy to account for; general results derived for first-generation TDI variables have valid counterparts in later generations. The modern tools for simulating the LISA data streams, such as `Synthetic LISA` (Vallisneri, 2005) and the `LISA Simulator` (Rubbo et al., 2003)), are automatically building these higher-order TDI signals into their architecture.

A.2. TDI Applications

While TDI is an important part of the overall LISA experiment, it also provides an unprecedented flexibility in the ability to analyze and calibrate the LISA data. TDI has many interesting applications which make it an invaluable tool for gravitational-wave astronomy.

As the raw data is subjected to pre-processing, there are several possible interferometric topologies that could be used (Estabrook et al., 2000), and which use as input between four and six links between the LISA sciencecraft. The ability to form a workable interferometric variable with less than six total links makes it possible for science analysis to proceed under a variety of contingencies. As with every TDI variable, each of these modes will yield a different instantaneous sky sensitivity. It has also been suggested that some limited science analysis is possible with only two links between two spacecraft, operating the observatory as a *xylophone detector* Tinto (1998).

One of the first results of a careful study of TDI is that the sensitivity curves for different TDI variables have different shapes across the LISA frequency band. Overall, the sensitivities are all very similar, but they can be reshaped to provide some flexibility. A particularly useful reshaping is the "Sagnac combination", usually denoted $\zeta(t)$, which has the interesting property that it is insensitive to gravitational waves at low frequencies. At frequencies below LISA's transfer frequency, $f_* \sim c/(2\pi L_o) \simeq 10\,\mathrm{mHz}$, the $\zeta(t)$ response to gravitational waves is suppressed by roughly a factor of 1000 (Armstrong et al., 1999; Estabrook et al., 2000). By contrast, instrumental noise is *not* suppressed. It has been suggested (Tinto et al., 2001) that this provides a method whereby the LISA instrumental noise can be assessed while the observatory is in orbit.

A deconvolution of the $\zeta(t)$ data channel should yield, at low frequencies, the pure instrumental noise uncontaminated by any but the strongest astrophysical signals. It has been suggested (Hogan and Bender, 2001) that a similar procedure can be used to determine how much of the observed low-frequency signal in a data stream is instrumental noise, and how much is simply confused astrophysical noise (*e. g.*, from the galactic population of compact binaries, or a stochastic cosmic gravitational-wave background, *cf.* section 6).

Another important TDI application related to the $\zeta(t)$ signal is the *zero signal solution* ZSS (Tinto and Larson, 2004, 2005). The ZSS can be used to determine the sky location of gravitational-wave sources without exploiting the modulation produced by LISA's yearly motion around the Sun. The technique takes a linear combination of a TDI triad to create a function characterized by two parameters, namely the sky location angles $\{\theta, \phi\}$. When a gravitational-wave signal is present in the data, it can be "zeroed out" by searching for the values of $\{\theta, \phi\}$ which minimize the ZSS. In essence, the technique uses the time of flight of a gravitational-wave signal across the LISA constellation to construct a triangulation of the source location. At high frequencies, a signal can be precisely zeroed out (to the level of the instrumental noise). At low frequencies, gravitational-wave signals cannot be zeroed out; they can only be suppressed in a fashion similar to $\zeta(t)$. This low-frequency behavior can be understood using the time-of-flight view of how the ZSS works; at low frequencies, the gravitational wavelength is larger than the entire LISA constellation, and so it is impossible to determine the direction that a gravitational wave is coming from, because the observatory is contained inside a single cycle of the wave, which is slowly varying on timescales long compared to the time of flight down the interferometer arms. Because it does not employ the Doppler modulation generated by LISA's motion around the Sun, the ZSS is particularly useful for burst localization on the sky.

An extensive review of the fundamentals of TDI may be found in Dhurandhar and Tinto (2005).

A.3. References

J. W. Armstrong, F. B. Estabrook, and M. Tinto. Time-Delay Interferometry for Space-based Gravitational Wave Searches. *ApJ*, 527:814–826, December 1999.

N. J. Cornish and R. W. Hellings. The effects of orbital motion on LISA time delay interferometry. *Class. Quantum Grav.*, 20:4851–4860, November 21 2003, also as arXiv:gr-qc/0306096.

N. J. Cornish and S. L. Larson. Space missions to detect the cosmic gravitational-wave background . *Class. Quantum Grav.*, 18:3473–3495, September 2001.

N. J. Cornish and L. J. Rubbo. LISA response function. *Phys. Rev. D*, 67:022001, January 15 2003.

R. J. Cruz, J. I. Thorpe, A. Preston, R. Delgadillo, M. Hartman, et al. The LISA benchtop simulator at the University of Florida. *Class. Quantum Grav.*, 23:751–+, October 2006.

S. V. Dhurandhar and M. Tinto. Time-Delay Interferometry. *Living Reviews in Relativity*, 8:4–+, July 2005.

F. B. Estabrook, M. Tinto, and J. W. Armstrong. Time-delay analysis of LISA gravitational wave data: Elimination of spacecraft motion effects. *Phys. Rev. D*, 62:042002, August 15 2000.

R. W. Hellings. Data processing for LISA's laser interferometer tracking system (LITS) . *Class. Quantum Grav.*, 18:4075–4080, October 7 2001a, also as arXiv:gr-qc/0012034.

R. W. Hellings. Elimination of clock jitter noise in spaceborne laser interferometers. *Phys. Rev. D*, 64(2):022002–+, July 2001b, also as arXiv:gr-qc/0012013.

C. J. Hogan and P. L. Bender. Estimating stochastic gravitational wave backgrounds with the Sagnac calibration. *Phys. Rev. D*, 64(6):062002–+, September 2001, also as arXiv:astro-ph/0104266.

S. L. Larson, W. A. Hiscock, and R. W. Hellings. Sensitivity curves for spaceborne gravitational wave interferometers. *Phys. Rev. D*, 62(6):062001–+, September 2000, also as arXiv:gr-qc/9909080.

S. L. Larson, R. W. Hellings, and W. A. Hiscock. Unequal arm space-borne gravitational wave detectors. *Phys. Rev. D*, 66(6):062001, September 2002, also as arXiv:gr-qc/0206081.

A. Povoleri and S. Kemble. LISA Orbits. In S. M. Merkowitz and J. C. Livas, editors, *LASER INTERFEROMETER SPACE ANTENNA: 6th International LISA Symposium*, volume 873, pages 702–706. AIP, 2006.

T. A. Prince, M. Tinto, S. L. Larson, and J. W. Armstrong. LISA optimal sensitivity. *Phys. Rev. D*, 66:122002–+, December 15 2002.

J. D. Romano and G. Woan. Principal component analysis for LISA: The time delay interferometry connection. *Phys. Rev. D*, 73:102001–+, May 2006.

L. J. Rubbo, N. J. Cornish, and O. Poujade. The LISA Simulator. In J. M. Centrella and S. Barnes, editors, *The Astrophysics of Gravitational Wave Sources*, volume 686, pages 307–310, October 2003.

L. J. Rubbo, N. J. Cornish, and O. Poujade. Forward modeling of space-borne gravitational wave detectors. *Phys. Rev. D*, 69(8):082003, April 2004.

D. A. Shaddock. Operating LISA as a Sagnac interferometer. *Phys. Rev. D*, 69: 022001, January 2004, also as arXiv:gr-qc/0306125.

D. A. Shaddock, M. Tinto, F. B. Estabrook, and J. W. Armstrong. Data combinations accounting for LISA spacecraft motion. *Phys. Rev. D*, 68:061303–+, September 15 2003, also as arXiv:gr-qc/0307080.

J. Sylvestre. Simulations of laser locking to a LISA arm. *Phys. Rev. D*, 70:102002, November 2004.

J. I. Thorpe, R. J. Cruz, S. R. Sankar, G. Mueller, and P. McNamara. First step toward a benchtop model of the Laser Interferometer Space Antenna. *Optics Letters*, 29:2843–2845, December 15 2004.

J. I. Thorpe, R. J. Cruz, S. Sankar, and G. Mueller. Electronic phase delay - a first step towards a bench-top model of LISA. *Class. Quantum Grav.*, 22:S227–S234, May 21 2005.

M. Tinto. LISA as a Xylophone Interferometer Detector of Gravitational Radiation. In W. M. Folkner, editor, *The Second International LISA symposium on the Detection and Observation of Gravitational Waves in Space*, volume 456, pages 121–127. AIP, 1998.

M. Tinto and S. L. Larson. LISA time-delay interferometry zero-signal solution: Geometrical properties. *Phys. Rev. D*, 70:062002, September 15 2004.

M. Tinto and S. L. Larson. The LISA zero-signal solution. *Class. Quantum Grav.*, 22: S531–S535, May 21 2005.

M. Tinto, J. W. Armstrong, and F. B. Estabrook. Discriminating a gravitational wave background from instrumental noise in the LISA detector. *Phys. Rev. D*, 63: 4081–4086, January 15 2001.

M. Tinto, D. A. Shaddock, J. Sylvestre, and J. W. Armstrong. Implementation of time-delay interferometry for LISA. *Phys. Rev. D*, 67:122003–+, June 15 2003.

M. Tinto, F. B. Estabrook, and J. W. Armstrong. Time delay interferometry with moving spacecraft arrays. *Phys. Rev. D*, 69(8):082001–+, April 2004.

M. Tinto, M. Vallisneri, and J. W. Armstrong. Time-delay interferometric ranging for space-borne gravitational-wave detectors. *Phys. Rev. D*, 71(4):041101–+, February 2005, also as arXiv:gr-qc/0410122.

M. Vallisneri. Geometric time delay interferometry. *Phys. Rev. D*, 72(4):042003–+, August 2005, also as arXiv:gr-qc/0504145.

M. Vallisneri. Synthetic LISA: Simulating time delay interferometry in a model LISA. *Phys. Rev. D*, 71:022001–+, January 2005.

B. LISA science objectives and scientific investigations

1. Understand the formation of massive black holes
   1.1 Search for a population of seed black holes at early epochs.
   2.2 Search for remnants of the first (Pop III) stars through observation of intermediate mass black hole captures, also at later epochs.
2. Trace the growth and merger history of massive black holes and their host galaxies
   2.1 Determine the relative importance of different black hole growth mechanisms as a function of redshift.
   2.2 Determine the merger history of 10^4 to $3 \times 10^5\, M_\odot$ black holes before the era of the earliest known quasars ($z \sim 6$).
   2.3 Determine the merger history of 3×10^5 to $10^7\, M_\odot$ black holes at later epochs ($z < 6$).
3. Explore stellar populations and dynamics in galactic nuclei
   3.1 Characterize the immediate environment of MBHs in $z < 1$ galactic nuclei from EMRI capture signals.
   3.2 Study intermediate-mass black holes from their capture signals.
   3.3 Improve our understanding of stars and gas in the vicinity of Galactic black holes using coordinated gravitational and electromagnetic observations.
4. Survey compact stellar-mass binaries and study the structure of the Galaxy
   4.1 Elucidate the formation and evolution of Galactic stellar-mass binaries: constrain the diffuse extragalactic foreground.
   4.2 Determine the spatial distribution of stellar mass binaries in the Milky Way and environs.
   4.3 Improve our understanding of white dwarfs, their masses, and their interactions in binaries and enable combined gravitational and electromagnetic observations.
5. Confront General Relativity with observations

   5.1 Detect gravitational waves directly and measure their properties precisely.
   5.2 Test whether the central massive objects in galactic nuclei are the black holes of general relativity.
   5.3 Make precision tests of dynamical strong-field gravity.

6. Probe new physics and cosmology with gravitational waves
   6.1 Study cosmic expansion history, geometry and dark energy using precise gravitationally calibrated distances in cases where redshifts are measured.
   6.2 Measure the spectrum of, or set bounds on, cosmological backgrounds.
   6.3 Search for burst events from cosmic string cusps.

7. Search for unforeseen sources of gravitational waves

C. Acronyms

A

AGN Active Galactic Nuclei

AMIGOS Analysis Methods for Interferometric Gravitational-wave Observations from Space

AO Analysis Object

B

BBH Binary Black Hole

BAM Blocked Annealed Metropolis-Hastings

C

CMB Cosmic Microwave Background

D

DAST Data Analysis Steering Team

DSN Deep Space Network

DFACS Drag Free and Attitude Control System

E

EOB Effective One-Body

ESA European Space Agency

EMP Experiment Master Plan

EMRI Extreme Mass Ratio Inspiral

G

GR General Relativity

GRS Gravitational Reference Sensor (synonym for LTP documents)

GW Gravitational Wave

I

IMS Interferometric Measurement System

IMRI Intermediate Mass-Ratio Inspiral

L

LSC LIGO Scientific Collaboration

LIST LISA International Science Team

LPF LISA Pathfinder

LSDC LISA Science Data Centre

LTP LISA Technology Package

LMB Low Mass Binaries

M

MCMC Markov Chain Monte Carlo

MBH Massive Black Hole

MBHB Massive Black Hole Binary

MoU Memorandum of Understanding

MHMC Metropolis-Hastings Monte Carlo

MOC Mission Operations Centre

MDC Mock Data Challenge

MLDC Mock LISA Data Challenge

N

NASA National Aeronautics and Space Administration

P

PN post-Newtonian

S

SNR Signal to Noise Ratio

SMBH Super-massive Black Hole

T

TDI Time Delay Interferometry

TBD to be defined

Z

ZSS zero signal solution