Analysis Methods for Interferometric Gravitational-wave Observations from Space (AMIGOS):
NASA Development Plan for the LISA Mission

Draft 0.2 – Dec 9, 2005

Tom Prince
thomas.prince@jpl.nasa.gov

Bonny Schumaker
bonny.schumaker@jpl.nasa.gov

Michele Vallisneri
michele.vallisneri@jpl.nasa.gov

…with the assistance and input of the LISA Mission Science Office at JPL
and of all the participants to the AMIGOS Workshop in Pasadena, CA, Oct 13-15, 2005.

Table of Contents

1. Introduction

The AMIGOS-LISA Development Plan is one of four formal documents associated with the LISA Project's Mission Science Office. [The others are the Science Requirements Document (SRD), the Data Management Plan (DMP), and the Science Management Plan (SMP).] Produced by the U.S. LISA Mission Science Office, the AMIGOS-LISA Development Plan is part of a living umbrella document that will include both development and implementation plans for LISA science data analysis. This broad area of research and development is herein referred to by the acronym "AMIGOS," for "Analysis Methods for Interferometric Gravitational-wave Observations from Space."

AMIGOS-LISA describes all U.S. LISA Project-funded activities related to the design, development, and implementation of methods, tools, and other resources for the scientific analysis of LISA data. The purpose of the AMIGOS-LISA Development Plan is to identify the key technical challenges in science data analysis for the LISA mission, and to recommend a development plan to address them. The AMIGOS-LISA Development Plan will later be merged into a joint NASA–ESA plan for data analysis.

The AMIGOS-LISA Development Plan complements NASA's LISA Technology Development Plan (which identifies key technical challenges in areas such as flight system hardware and software, and which recommends a development plan to address them), and its evolution toward a joint ESA-NASA data analysis plan is analogous to that of the Technology Development Plans held by the NASA and ESA LISA Project Offices.

Some AMIGOS activities are now included explicitly in the Product Breakdown Structure for the LISA Project's Mission Science element; others will become so after refinement through less formal investigative efforts, in the usual manner of scientific insights and discoveries. Some of the research and development needed to utilize fully the LISA scientific data, such as efforts in related areas of general relativity and astrophysics, will be carried out through support from sources other than the LISA Project. This document does not attempt to address all of this work, but it does attempt to describe generally what efforts are needed and more specifically what efforts may be carried out with funding obtained directly from the LISA Project.

How to read this document

Previous drafts of this documents (labeled v0.1x) resulted from extensive efforts, throughout 2005, to collect inputs on data analysis planning from the LISA Project and the gravitational-wave community; these efforts culminated in the 2.5-day AMIGOS workshop, held in Pasadena on Oct. 13–15 and attended by more than 30 gravitational-wave–science experts (see www.tapir.caltech.edu/amigos for the full "live" proceedings of the workshop).

This draft of the document (v0.2) represents the authors' attempt to select, organize, and prioritize the proposed development areas and tasks, drawing a reasoned path for LISA data analysis through the initial phases of the mission. The core of the planning effort are the broad *goals* given in Sec. 3.1, which define the desired outcomes of planning either in functional terms (e.g., "demonstrate a data analysis system than can achieve the minimum mission science measurement requirements") or in terms of deliverables crucial to other project elements (e.g., "determine constraints on posed on mission design by the data-analysis architecture"). The goals are supported by *high-level milestones* (Sec. 3.2), which are classified according to R&D area: Instrument Performance and Characterization (IPC), Computational Infrastructure (COMP), Analysis Methods (MET, subdivided in Source-Specific, MET-SS, and Global, MET-GF), Astrophysics (AST, subdivided in Waveforms and Models, AST-WF, and Event Rates, Populations, and Inverse Problems, AST-RT), and General Relativity (GR). Section 3.3 (currently missing) focuses on risk management. Section 4 examines the high-level milestones, breaking them off into prioritized intermediate milestones, which are possible objectives for targeted contracts or in-house research efforts. Estimated effort levels are given for the intermediate milestones proposed for Phase A.

Section 2 provides background on the LISA Data Analysis problem, with Sec. 2.1 concentrating on its novel and challenging aspects, Sec. 2.2 giving a brief overview of LISA sources with the corresponding Science Measurement Requirements, and Sec. 2.3 (currently missing) giving a high-level overview of the possible architecture for LISA data analysis. Section 5 will later contain the management plan for AMIGOS activities.

2. Background

2.1 The character of LISA data analysis

Although experience from searching for gravitational waves with ground-based detectors will be valuable to the development of LISA data-analysis capabilities, there are fundamental differences between LISA and its ground-based counterparts, in both the instruments and the GW sources, that present new difficulties for data analysis and require novel targeted solutions. These differences include:

Detector response—Over much of the LISA measurement band, the physical size of LISA is comparable to or barely smaller than the wavelength of the gravitational waves. LISA's response therefore depends sensitively on the light-travel time across the array. (By contrast, ground-based detectors operate in the "long-wavelength" limit.)

Source confusion—In the middle-to-lower part of the LISA measurement band, the LISA data output will be dominated by a diffuse foreground of white-dwarf binaries from our Galaxy. Ground-based detectors, on the other hand, are dominated throughout their measurement band by instrument noise.

Continuous sources—Most sources for ground-based detectors are short-lived. In contrast, most LISA sources will be long-lived and visible to LISA for weeks, and usually for much longer. (A notable exception is possible bursts from string cusps or kinks.)

Perhaps the major challenge for LISA data analysis is that the time series will be signal-rich and most of the GWs will be "on" simultaneously. This means that construction of a signal catalog is likely to require the simultaneous identification of most or all the GW signals present in the data set. The best fit will probably be obtained iteratively, but algorithms do not yet exist to do so—nor is there current theoretical understanding of how well we can expect to do. This is a significant data-analysis challenge (referred to in this document as the *global-fit problem*) that will require efforts from the larger GW science community as well as within the LISA project.

2.2 An overview of LISA science goals and requirements by source class

In this section we briefly summarize the data analysis requirements that flow from scientific and technical considerations. The high-level science requirements for LISA are stated in the LISA Science Requirements Document (SRD), organized according to the nature of the different sources of GWs to be observed. For each category of sources, the SRD lists both observational requirements (S) and measurements requirements (M). One additional source category is included here, that of unanticipated GW sources; however, its implications for analysis methods are addressed in Section 4.1 since its detection depends entirely on characterization of instrument science operation and performance.

2.2.1 Compact binaries in the Milky Way and nearby galaxies

The study of galactic binaries by LISA will provide a rich yield of new information on compact binaries, including a complete three-dimensional map of the ultracompact binaries in the Milky Way and nearby galaxies, and a much improved knowledge of the evolutionary pathways of compact binaries. In addition, only few months of LISA data will be needed for the positive detection of the nearest and strongest known binaries: these sources (*verification binaries*) will play an important role in the initial verification of LISA's instrument performance.

It is useful to classify compact-binary systems according to the character of their prospective detection.

Verification binaries—These are sources that have been identified in electromagnetic observations, so their sky positions and orbital periods are known well, and their masses, distances, and orbital inclinations are known reasonably well. The subset of systems evolving primarily through GW radiation (without significant mass transfer) can function as reliable calibrators for the LISA sensitivity, because they emit gravitational waveforms that are described accurately by slow-motion, weak-field general relativity. Additional hydrodynamical modeling is needed to reliably detect mass-transferring systems, and to characterize their properties.

Unknown but resolvable galactic binaries—At frequencies higher than 2–3 mHz, the LISA output contains enough information to resolve, in principle, $\sim 10^4$ distinct binaries in the Galaxy. This is however a very challenging task, because it entails the *simultaneous* detection and characterization of thousands of individual sources that are present at all times in the LISA science data stream, and that overlap significantly in frequency space.

Unresolvable binaries—At frequencies lower than 2–3 mHz, binaries become so numerous that they are not resolvable individually, except for the nearest and strongest systems. The science goal then shifts to constraining the overall amplitude and the first few angular moments of this diffuse background (see Sec. 2.2.4).

The basic data-analysis requirement for the study of compact binaries by LISA is the capability to detect individual binaries and to determine their orbital periods, orbital-period derivatives, orbital inclinations, chirp masses, and sky positions in the presence of instrument noise and of other GW sources (including other resolvable binaries, the unresolvable binary backgrounds, and sources of other classes).

Relevant SRD requirements

S1.1, S1.2, S1.3, M1.3, M1.4, S2.1, M2.3, M2.4, M2.5

2.2.2 Massive black-hole (MBH) binary mergers

These sources include *intermediate black-hole* (IMBH) binaries, defined as systems where the primary has mass smaller than $\sim 3\times10^4$ Solar masses, and *supermassive black-hole* (SMBH) *binaries*, defined as systems where the primary has mass greater than 3×10^4 Solar masses, and the mass ratio is larger than 0.01 (systems with smaller mass ratios are classified as *extreme–mass-ratio inspirals*, EMRIs, and are discussed in Sec. 2.3).

These mergers will generate strong signals even if they occur at high redshifts, but event rates are very uncertain: current estimates range between a few and a few hundred detectable mergers over the LISA lifetime. Merger waveforms will encode detailed information about the physical parameters of the system and about strong-gravity general-relativistic dynamics.

The basic data-analysis requirement for the study of massive black-hole binaries by LISA is the capability to detect merging SMBH and IMBH systems and to determine their parameters, including component masses, orbital periods and inclinations, and sky positions. In addition, in order to enable concurrent electromagnetic observations of the mergers, the data-analysis system must be capable of detecting MBH systems in the early part of their inspiral, and of accurately determining their parameters (especially sky position, luminosity distance, and the predicted time of merger). Because MBH binary mergers might be the strongest sources in the LISA data stream (possibly at any given time, if their event rates are large), it is crucial to model their waveforms accurately enough that they don't interfere with the detection and characterization of weaker sources in the same frequency band.

Relevant SRD Requirements

S3.1, S3.2, S3.3, S3.4, M3.5, M3.6, S4.1, M4.3, M4.4 (SRD v2.7 050118)

2.2.3 Extreme–Mass-Ratio Inspirals (EMRIs)

The detection of the inspirals of compact objects into the supermassive objects at the centers of galaxies (EMRIs) may provide some of the most exciting scientific payoffs in the LISA mission.

From a *general-relativistic viewpoint*, the waveforms will encode the spacetime geometry induced by the central object, allowing for high-precision measurements of its multipole moments, which could confirm black-hole no-hair theorems, or possibly identify the central object as something other than a black hole (e.g., a solitonic star or a naked singularity). The waveforms could also be used to measure the response of the central body to the tidal gravity of the orbiting object, again confirming or disproving the predictions of general relativity.

From an *astrophysical viewpoint*, a catalog of detected EMRI events would probe the astrophysics of the dense clusters in galactic nuclei, including the mass demographics of cluster objects and the presence of intermediate-mass black holes in the nuclei and dense disks around the central object.

The SRD states the LISA EMRI requirement as the "capability to detect the gravitational waves emitted during the last year of inspiral for a 10-solar-mass black hole orbiting a 3×10^5–3×10^6 solar-mass black hole at 1 Gpc, with an optimal SNR of 40 or better," with the "ability to determine the evolution of signal polarization due to precession of the orbital plane" (as required for the general-relativistic tests). According to Gair et al. (2004), this capability would allow for the detection of hundreds of EMRI events under plausible assumptions on capture event rates.

The primary data-analysis challenge for the successful *detection* of EMRI events is overcoming the sheer complexity of the expected waveforms, which have a rich spectral structure and depend sensitively on several source parameters. The challenge is compounded by the necessity to build a variety of non-black-hole and non-general-relativistic features into the waveforms. The theoretical waveforms used for detection need not necessarily be faithful representations of the true physical waves, but they need to encompass a large enough variety of signal behaviors to recover most of the available signal power while avoiding unnecessary noise-induced false alarms.

Preliminary studies (Gair et al. 2004, Barack and Cutler 2004a) suggest that the *coherent* (i.e., matched-filtering) detection of EMRI waveforms with integration periods of a year or more is computationally unfeasible, because of the number of separate source parameter sets that would need to be tested to maintain phase coherence between the physical signal and the test (*template*) waveforms. The practical alternative is a *partially coherent* scheme, whereby the LISA output is filtered coherently over periods of few weeks, after which the strongest triggers are pieced together to reconstruct the waveform in its entirety. This scheme is analogous to the *stack-slide* (and Hough-transform) methods used to search for continuous pulsar signals using ground-based gravitational-wave detectors; but it is complicated by the multi-dimensional character of the *stacking* step.

The primary data-analysis challenge for the successful *interpretation* of EMRI events is computing high-accuracy waveforms that can be used to estimate the source parameters reliably. Both the detection and interpretation steps must take into account the presence of foreground and background signals from other sources, with source parameters modeled only to power- and confusion-limited precision.

Relevant SRD Requirements

S5.1, S5.2, M5.1, M5.2, M5.3, M5.4, M5.5, M5.6

2.2.4a Diffuse galactic background

The GW *diffuse galactic background* (DGB) is a double-edged sword for LISA: it is itself an interesting astrophysical source of GWs, but it is also a source of quasistochastic noise that can overwhelm weaker sources and confuse the signals from stronger sources. (For most LISA GW sources it is technically a "foreground" because its sources are physically closer to us.) In order for data-analysis algorithms to account for this quasistochastic noise, it must itself be characterized. For example, the LISA data will be used to infer the frequencies and large-scale angular distribution of the DGB. This information permits inferences about the spatial distribution and evolution of compact binary systems in our galaxy. Further, comparison of the DGB angular distribution with the three-dimensional map of individually-resolved compact provides gives information about the "demographics" of the DGB. Finally, knowledge of the DGB angular distribution may permit recognition of and inferences about any isotropic diffuse background of extragalactic or cosmological origin.

The basic observational requirement is to determine the overall amplitude and the first few spatial moments of the spatial distribution of the diffuse galactic background. This will require characterizing instrument noise well enough to distinguish it from the DGB, which can be accomplished by using specific TDI observables that are insensitive to gravitational waves (such as the symmetrized-Sagnac observable "zeta"), and by modeling as-flown instrument noise as discussed in Sec. 4.1. It is also important to understand how the presence of the background affects the search for other signals, and the determination of their parameters.

Relevant diffuse galactic background SRD Requirements

S2.2, M2.3, M2.5

2.2.4b Diffuse cosmic background

Gravitational-wave observations of the *diffuse cosmic background* can be broken into two broad goals: detection and characterization. The diffuse background might be composed of many components, including:

– a Cosmic Gravitational-wave Background (CGB) of waves originating from an era between the Big Bang and recombination (analogous to the Cosmic Microwave Background);

– a confusion-limited background of extra-galactic binaries (EGB) (Farmer and Phinney 2003);

– a background of weak gravitational waves from interacting massive black-hole binaries that are still far from coalescence (these binaries can spend 10^2 to 10^3 years radiating in the LISA band, so if the observed coalescence rate is 1 to 10 per year, there will always be 10^2 to 10^5 massive BH binaries "on" in the LISA data);

– a foreground of weak EMRI signals (Barack and Cutler 2004b).

The primary observational requirement is either to detect an isotropic gravitational-wave background (which would be a fundamental discovery) or to establish upper limits on its strength, which requires the capability of distinguishing instrument noise from an isotropic GW background. At low frequencies, this should be possible by using GW-insensitive TDI observables; at higher frequencies, instrument noise would probably need to be modeled independently of the TDI data.

Relevant diffuse cosmic background SRD Requirements

S6.1, S6.2, M6.1, M6.2 (SRD v2.7)

2.2.5 Unanticipated GW sources

Historically when new observational windows for astronomy are opened the strongest or most interesting sources were not anticipated [*e.g.,* Kellermann and Sheats 1983, Thorne 1987, Phinney 2002]. There is currently no science requirement on the capability to detect unexpected GW sources. However, this capability is entirely dependent on the accuracy with which instrument performance is understood and characterized, so we refer the reader to Sec. 4.1.

2.3 A high-level overview of the LISA data-analysis architecture

…

3. LISA Data-Analysis Development Planning

3.1 LISA Data-Analysis Goals

Table 3.1, below, summarizes the broad goals in the AMIGOS-LISA Development Plan. The goals track the science and technology readiness of the LISA Data-Analysis System; their completion is planned before standard project reviews (SRR, PDR, LRR), whose calendar date depends on the overall project schedule.

The succession of goals was designed according to the following criteria:
- priority was given to developing and demonstrating the capability to meet the minimum and baseline science measurement requirements;
- priority was given to exploring the requirements posed by the data-analysis system on the design of mission systems and operations;
- priority was given to determining the science-data–based procedures necessary for on-orbit commissioning and instrument verification.

Each goal is supported by a set of high-level milestones, described in section 3.2.

Phase A	1. Demonstrate *proof-of-concept* data-analysis system that shows: – *demonstrated ability* to meet the *minimum* mission science measurement requirements, including detection and initial parameter estimation, as evidenced by simulations on *realistic* data sets; – *capability* to achieve the *full* baseline-mission science measurement requirements, including detection, initial parameter estimation, and global-fit procedures for multiple source classes, as evidenced by analytic arguments and simulations on representative data sets. 2. Determine the constraints posed on mission design by the science requirements and by the candidate data-analysis architecture. 3. Develop plan for on-orbit science commissioning and initial instrument performance verification.
Phase B	1. Demonstrate *prototype* data-analysis system that shows capability of achieving the *full* baseline-mission science measurement requirements, as evidenced by extensive simulations on *high-fidelity* data sets that include real-mission features such as data gaps and glitches. [Where accurate source waveforms are not mature, demonstrations would be run with “kludge” waveforms of realistic complexity.] 2. Finalize data-analysis requirements on mission design. 3. Demonstrate prototype system for on-orbit commissioning and initial instrument performance verification. 4. Complete assessment of computational requirements of all data-analysis applications (searches, global fits, parameter-estimation studies).
Phase C–D	1. Deliver formal specification of *production* data-analysis system, including extensive description and documentation of computational infrastructure and data-analysis applications (suitable for formal verification when possible). 2. Deploy data-analysis system. 3. Demonstrate its capability to process six months of LISA data in real time, with a full suite of data-analysis applications. 4. Plan for and secure adequate computational resources.

Table 3.1. AMIGOS-LISA Development Plan Goals by Mission Phase

3.2 High-level milestones

Table 5.1, below, summarizes the high-level milestones in the AMIGOS-LISA Development Plan. Milestones are organized by development area (e.g., *Instrument Performance and Characterization*, *Computational Infrastructure*), and they are articulated in three phases, corresponding to the gates of Sec. 3.1.

Section 4 provides background and technical details for each high-level milestone, including a breakdown into intermediate milestones with priority ratings and estimated effort levels.

label	*description*	*rationale*
IPC (Sec. 4.1)	**Instrument Performance and Characterization**	
IPC-A	Explore mission design requirements from data analysis. Plan on-orbit science commissioning and instrument verification procedures. Demonstrate proof-of-concept tools for on-orbit noise characterization.	Retires risk and expense of modifying mission design at later stage to accommodate data-analysis requirements. Prioritized to allow sufficient lead time for definition of initial mission operations. Needed to enable development of data-analysis applications that will work with as-measured noise.
IPC-B	Finalize mission design requirements from data analysis. Deliver prototype on-orbit science commissioning and instrument verification tools. Increase fidelity of noise modeling and simulation. Develop and test robust noise-characterization tools.	Needed for final mission design. Prioritized to allow sufficient lead time for definition of initial mission operations. Needed to develop and test robust data-analysis applications that will work with as-measured noise.
IPC-C	Demonstrate and document high-fidelity model of instrument noises and performance. Deliver and test production system for noise characterization.	
COMP (4.2)	**Computational-Infrastructure**	
COMP-A	Deploy and demonstrate *initial computational infrastructure* testbed for use as testground in data-analysis application development.	Needed to kickstart development of data-analysis applications and to demonstrate proofs of concept with simulations on realistic data sets.
COMP-B	Deploy and demonstrate *prototype computational infrastructure* for extensive testing of serial and parallel data-analysis algorithms and pipelines and for accurate characterization of computational requirements. Assess computational requirements for infrastructure.	Needed to support and test robust data-analysis applications on high-fidelity data sets. Prioritized to allow lead time toward planning for and securing computational resources.
COMP-C	Deliver specifications for *final production computational* infrastructure. Deploy final computational infrastructure and secure computational resources. Prepare for transition to production infrastructure interfaced with data center.	

Table 3.2. AMIGOS-LISA Development High-Level Milestones (continues overleaf…)

MET-SS (4.3a)	**Source-Specific Detection and Parameter-Estimation Methods and Tools**	
MET-SS-A	Develop and demonstrate (by extensive simulations on realistic data sets) *proof-of-concept* methods for reliable detection and initial parameter estimation for the sources included in the *minimum* science measurement requirements. Develop and demonstrate (by analytical arguments and simulations on representative data sets) *proof-of-concept* methods for reliable detection, parameter estimation, and concurrent fitting for all sources included in the *full* baseline-mission science measurement requirements.	Needed to demonstrate capability to meet minimum science measurement requirements. Prioritized to allow lead time toward demonstrating capability to meet baseline-mission science measurement requirements.
MET-SS-B	Develop and demonstrate (by extensive simulations on high-fidelity data sets including realistically complex waveforms and source populations of realistic magnitude) robust *prototype* methods for reliable detection, accurate parameter estimation, and concurrent fitting for all sources included in the *full* baseline-mission science measurement requirements. Assess computational requirements for source-specific data-analysis applications and global-fit plugins.	Needed to demonstrate capability to meet baseline-mission science measurement requirements. Prioritized to allow lead time toward planning for and securing computational resources.
MET-SS-C	Deliver specifications for detection and parameter-estimation stand-alone applications and global-fit plugins. Deploy production source-specific applications and plugins.	
MET-GF (4.3b)	**Global-Fit Procedures**	
MET-GF-A	Investigate, develop, and demonstrate (by simulations on representative data sets) alternative *candidate* procedures for the *global-fit* detection and parameter estimation of overlapping signals from different source classes.	Prioritized to allow lead time toward demonstrating capability to meet baseline-mission science measurement requirements. [*The concurrent-detection problem for LISA has no proven analog in any other mission or experiment.*]
MET-GF-B	Select, develop and demonstrate (by extensive simulations on high-fidelity data sets) robust *prototype* global-fit procedures that can be scaled up to source populations of real magnitude. Assess computational requirements for global-fit pipelines (using source-specific information from MET-SS-B).	Needed to demonstrate capability to meet baseline-mission science measurement requirements. Prioritized to allow lead time toward planning for and securing computational resources.
MET-GF-C	Finalize analysis architecture and deliver specifications for production global-fit procedures. Deploy production global-fit procedures.	
AST-WF (4.4a)	**Astrophysical Waveforms and Models**	
AST-WF-A	Track progress in community-based research on astrophysical models needed to secure maximum science payoff from LISA.	Needed to inform development of source-specific detection and parameter-estimation tools.
AST-WF-B	Review, encourage, and commission community-based research on astrophysical models needed to secure maximum science payoff from LISA. in crucial areas where early lead-in is needed to secure maximum science payoff from LISA.	Needed to inform development of source-specific detection and parameter-estimation tools; prioritized to allow for lead time and contingency in theoretical research. [*Can't legislate theoretical progress*].
AST-WF-C	Compile complete set of astrophysical models to allow the optimal extraction of scientific information from the LISA data.	

Table 3.2. (…continued.) AMIGOS-LISA Development High-Level Milestones (continues overleaf…)

AST-RT (4.4b)		**Astrophysical event rates, populations, and inverse problems**	
	AST-RT-A	Characterize astrophysical event rates and populations as needed to finalize science measurement requirements, to prepare for on-orbit commissioning, and to inform the development of data-analysis techniques.	Retires risk and expense of over- or under-designing mission; prioritized to allow sufficient lead time for definition of initial mission operations; needed to inform development of source-specific detection and parameter-estimation tools.
	AST-RT-B	Increase fidelity of event-rate and population models.	Needed for final mission design and to inform development of source-specific detection and parameter-estimation tools.
		Formulate inverse problems to obtain the maximum scientific payoff from the LISA data.	Prioritized to allow for lead time in development and testing of data-analysis applications to solve inverse problems.
	AST-RT-C	Extend AST-RT-B work.	
GR (4.5)		**General relativistic waveforms and models**	
	GR-A	Investigate general-relativistic waveforms and models, as needed to finalize measurement requirements, to inform the development of data-analysis techniques, and to indicate need for community-based effort where major advances are needed.	Retires risk and expense of over- or under-designing mission; needed to inform development of source-specific and global-fit detection and parameter-estimation tools; prioritized to allow for lead time and contingency in theoretical research.
	GR-B	Develop and deliver tools to efficiently compute *realistic* general-relativistic waveforms and models, as needed to develop and test source-specific and global-fit data-analysis applications.	Needed for final mission design and to inform development of source-specific detection and parameter-estimation tools; prioritized to allow for lead time and contingency in theoretical research
	GR-C	Develop and deliver tools to efficiently compute *high-fidelity* general-relativistic waveforms and models, as needed for production data-analysis system and to secure maximum science payoff (optimal detection rates, accurate parameter studies).	

Table 3.2. (…continued.) AMIGOS-LISA Development High-Level Milestones

3.3 Risk assessment and management

For data and science analysis methods, management of risk is accomplished chiefly by verification schemes, such as double-blind simulations in which known sources or known noise are injected into models of the LISA detector, and the results of analysis compared with the known inputs. Risk identification is the first challenge, after which risk assessment may be carried out with a standard matrix comparing Likelihood with Consequences or Impact. In this case, the Consequences are weighed against meeting the Science Requirements. The Likelihood is a combined assessment of the present sophistication and fidelity of the analysis methods, and the difficulty and time/budget involved in achieving the desired sophistication and fidelity.

The top five potential risks identified at this time are listed in Table 3.3 (higher number is worse).

Risk	**Likelihood (1-5)**	**Impact (1-5)**
1. …	…	…
2. …	…	…
3. …	…	…
4. …	…	…
5. …	…	…

Table 4.3. Top five potential risks for LISA data analysis.

These risks will be tracked and mitigated as the development program continues. None of them is expected to limit success in terms of meeting the LISA Science Requirements, but they will be used to guide the development program. A complete set of risks including mitigation strategies will be prepared in the future.

4. Technical approach

4.1 IPC – Characterization of instrument science operation and performance

4.1.1 Background

The probabilistic nature of gravitational-wave detection and characterization implies that a thorough understanding of instrument noise is needed to make reliable observations (e.g., to asses detection and false-alarm probabilities) and, indeed, to plan and implement detection strategies. This is especially true for sources without waveform models, such as stochastic backgrounds and some classes of bursts. History has shown that when a new observational window for astronomy is opened, the strongest and most interesting sources are those that were not anticipated (see, e.g., Kellermann and Sheats 1983, Thorne 1987, Phinney 2002). The LISA data analysis methodology should be capable of recognizing gravity waves from sources that were not anticipated.

To accomplish this, LISA science analysis procedures will require these interrelated capabilities:

- (IPC-3) a thorough signal-independent procedure to characterize *as-flown* instrument noises (including their levels, spectral shapes, degrees of nonstationarity, and deviations from Gaussian statistics);
- (IPC-4) a complete catalog of instrument noise transfer functions (for known, individual sources of noise and all the TDI data combinations), so that stretches of noise-only data could be positively identified as such. This should also include a processing scheme which treats simultaneous multiple TDI time series to isolate specific noise problems using different transfer functions;
- (IPC-5) a template-independent procedure to apply the polarization- and source-position-dependent GW signal transfer functions (for all TDI data combinations) to time series analysis so that an unanticipated waveform would be recognized.

In addition, this development area includes the investigation of the design constraints posed on the mission by data-analysis requirements and strategies (IPC-1), and the planning of the initial on-orbit science commissioning and instrument verification (IPC-2).

4.1.2 Intermediate milestones and development tasks

High-level milestones for Instrument Performance and Characterization Total estimated effort for Phase A: 2.6 FTE		Explore mission design requirements from data analysis. Plan on-orbit science commissioning and instrument verification procedures. Demonstrate proof-of-concept tools for on-orbit noise characterization.	Finalize mission design requirements from data analysis. Deliver prototype on-orbit science commissioning and instrument verification tools. Increase fidelity of noise modeling and simulation. Develop and test robust noise-characterization tools.	Demonstrate and document high-fidelity model of instrument noises and performance. Deliver and test production system for noise characterization.
Intermediate milestones		**Phase A**	**Phase B**	**Phase C–D**
IPC-1	Establish system-design requirements driven by data-analysis system Estimated effort for Phase A: 0.4 FTE Priority for Phase A: top	Identify trade studies on design tradeoffs.	Extend trade studies and make final recommendations.	

IPC-2	Design data-analysis procedures for on-orbit commissioning, including: - verification of instrument-science performance; - estimation of instrument noise; - removal of high-frequency Galactic binaries and estimation of low-frequency binary background Estimated effort for Phase A: 0.4 FTE Priority for Phase A: top	Draft architecture for on-orbit science commissioning. Examine use of verification binaries for system diagnostics.	Refine architecture. Deliver prototype data-analysis tools for on-orbit commissioning and test them with simulated data.	Deliver production capabilities, and demonstrate them in blind tests using simulated data.
IPC-3	Characterize as-flown instrument noises Estimated effort for Phase A: 1.0 FTE Priority for Phase A: middle/top	Develop tools to characterize level and shape of noise spectrum, in the presence of GW signals, by processing multiple TDI data combination time series. Develop tools to identify intervals of noise nonstationarity and trace these using transfer functions to the subsystem causing the nonstationarity. Explore effects of data gaps and glitches, and requirements to be set on them.	Develop techniques to identify nonlinearities in the noise-spectrum; exploit this to help with noise-and-signal identification using higher-order spectra. Deliver and test prototype capabilities with simulated model of low-level TDI reduction.	Deliver production capabilities, and demonstrate them prior to launch in blind tests using simulated data.
IPC-4	Catalog and Apply Instrument Noise Transfer Functions Estimated effort for Phase A: 0.4 FTE Priority for Phase A: middle	Develop transfer functions to the TDI time series for new noise sources as those sources are identified. Verify or refute that aggregate optical path noise (including e.g., pointing noise) has the same transfer function as shot noise.	Evaluate the extent of anticipated nonlinearities in how the fundamental noises enter the TDI laser-noise-canceled LISA data sets. Apply transfer functions to higher moments (e.g., bispectra) to evaluate what extra information can be obtained about nonlinearities in the system. Demonstrate spectral signatures of proof mass and optical path noises using simulated data. Feed results to COMP-3.	Extend Phase A and B work and feed results to IPC-2

IPC-5	Catalog and apply GW signal transfer functions Estimated effort for Phase A: 0.4 FTE Priority for Phase A: middle	Develop a template-independent procedure, processing multiple TDI data combinations, to classify science data intervals as probably “noise-like” or “candidate-signal-like”. Develop techniques exploiting the differing noise- and signal-couplings to the TDI data combinations to evaluate presence or absence of GW signals in specific intervals of data.	Extend Phase A work and feed results to MET-SS-5.	Demonstrate production capabilities and demonstrate them in blind tests using simulated data.

4.1.2 Status

4.1.3 Risk Reduction

4.2 COMP – Computational infrastructure for data analysis

4.2.1 Background

The case for early computational infrastructure development

The science analysis of the LISA data is a complex and layered endeavor: it will require the simultaneous detection of a large number of widely different signals, all entangled in the LISA data stream; it will deliver multiple data products, issued incrementally and periodically, such as the reduced TDI streams, various source catalogs, source population upper limits, and other science studies; and it will do all of this by involving multiple agents, both at the project and at the academic and research institutions contracted for data-analysis tasks.

At a minimum, the LISA computational infrastructure for data analysis must provide mechanisms for the storage, delivery, reduction, and conditioning of data; tools to produce simulated waveforms and noise; file format standards for the raw data and the science products; guidelines for the development of data-analysis codes; and access to computational resources. The data-analysis applications (such as the searches for specific sources) build on, and plug into the infrastructure.

The LISA computational infrastructure for data analysis can also play a unique strategic role among the research areas discussed in this document, in that:

- Early on, the infrastructure can provide a common testground for the agents developing data-analysis algorithms and applications, helping them to harmonize their efforts and to avoid duplication of effort.
- The development of common data formats and knowledge management functions provides an organizing principle for research and development in the simultaneous-detection aspect of the data-analysis problem.
- The verification of data-analysis milestones can be implemented pervasively with numerical experiments (such as mock data challenges) that take place within the infrastructure, providing a standard framework for the validation of externally contributed data-analysis applications by the project.
- The risk inherent in the transition from testing with simulated data to actual science operations can be reduced by incrementally incorporating realistic features (such as nonstationary noise, gaps, and the effects on data of housekeeping tasks) in the simulations. Ideally, the beginning of actual science operations would happen simply by switching off the simulators and plugging in the real instrument.

To achieve these goals, a prototype infrastructure must be deployed early in the implementation of the data-analysis plan, and strong waveform and noise simulation capabilities must be built in from the beginning.

It is understood that the initial data reduction (up to calibrated phase measurements and perhaps to TDI observables) is generally considered an instrument function that is covered separately in the technology development plan. However, a transparent understanding of the reduction pipeline is crucial to the accuracy and reliability of the data-analysis algorithms and of the science studies; it would be very dangerous to assume that the science studies can proceed naively from reduced data delivered “to specifications.” For these reasons, the computational infrastructure should evolve to include accurate models of low-level data reduction: this will allow the data-analysis applications to be tested in realistic conditions, and to incorporate essential insights from the data-reduction process. The work to develop the models of data reduction will have the added value of fostering a tight coupling between data analysts and instrumentalists, which is especially necessary because of the novel nature of gravitational-wave detection in space, and which has proved invaluable in ground-based experiments of comparable complexity, such as LIGO.

Likewise, other components of the LISA computational infrastructure (such as data storage) may also be considered as a separate project function that is implemented independently of data analysis. It is however appropriate to include such components in this document because targeted research and investments are needed to develop their specifications and their interface to data-analysis applications in the form that is most advantageous to the ultimate science payoff of the mission.

The components of the LISA computational infrastructure

Whereas the computational infrastructure for data analysis is an integrated system that draws strength from the synergistic interaction of its elements, it is still possible to identify challenges and requirements for several of its subsystems.

- *Data storage and delivery* (COMP-1). The data storage and delivery system will provide rapid and redundant access to the LISA data at multiple levels. According to the draft Science Data Management Plan, level 0 refers to onboard raw data, level 1A to onboard edited data (the phase measurements), level 1B to resampled data (the TDI streams), level 2 and greater to derived data products (such as source catalogs and source-foreground-free TDI data streams). If TDI is implemented onboard, level-1A-to-1B data reduction will be performed on the spacecraft, and only the level-1B dataset will be telemetered to Earth, except perhaps for brief periods in the commissioning phase; if TDI is implemented in postprocessing (Shaddock et al. 2004), the level-1A dataset will be telemetered to Earth, where level-1A-to-1B data reduction will be performed. The data storage and delivery system should be accessible using a lightweight, scalable, user-friendly library with bindings for multiple computer languages. The design of the library must balance ease of access with the enforcement of validation, consistency, and traceability checks.
- *Data reduction and conditioning* (COMP-2). The data reduction and conditioning library will include functions to generate higher-level from lower-level datasets (e.g., by subtracting foreground sources according to provisional data models; for postprocessed TDI, by computing TDI variables from the primitive phase measurements); it will include also other functions needed in data analysis, such as the generation of noise power spectral densities. The development of the library must proceed in concert with the characterization of the science operation and performance of the instrument. As for the data storage and delivery functions, the emphasis should be on providing a lightweight, scalable, user-friendly library with bindings for multiple computer languages; the design of the library must balance ease of access with the enforcement of validation, consistency, and traceability checks.
- *Source and noise simulators* (COMP-3). The source simulators will accurately reproduce the level-1B LISA instrument response to a wide range of gravity-wave sources, for input in the design and testing of data-analysis applications, and especially for the development of a simultaneous-detection paradigm. Although the source simulators will rely in part on the waveform-generation modules included in the data-analysis applications (e.g., as needed for signal template generation), they should also encompass a variety of unexpected or nonstandard signals, useful for testing the robustness of the data-analysis applications. The noise simulators will produce increasingly accurate and realistic realizations of the expected LISA noises, initially including all fundamental noises and later also technical noises, simulated telemetry and interferometry events that cause signal disruption or degradation, and so on. (The most accurate way to simulate the latter would be to create synthetic level-0 data and submit it to the data-reduction pipelines ultimately implemented for LISA.) The data streams produced by the simulators will be stored and accessed through the data storage and access libraries. Source and noise simulators will also play a role in the characterization of instrument science operation and performance.
- *File formats and data structures* (COMP-4). The LISA file formats should build on established standards (such as XML) to provide flexible solutions that can be used early on to represent the simulated data fed to the prototype data-analysis applications, and later grow to hold the actual instrument output. The formats should cover all the data levels, including level-2 and higher data models. These play an important role in the simultaneous-detection problem, as they encode the provisional and probabilistic knowledge about the sources present in the data stream. For these, the challenge will be to design and update the structure of data models as our understanding of simultaneous detections progresses. Lightweight and user-friendly input–output and parsing libraries for the LISA file formats should also be provided.
- *Data-analysis application guidelines* (COMP-5). The computational infrastructure should provide guidelines for the implementation of data-analysis applications, to ensure their interoperability with the infrastructure, and to facilitate code reuse between applications. Indeed, it would be desirable for application code to be publicly available to all the data-analysis agents, to encourage mutual scrutiny and verification. The guidelines can also help the project in validating, tracking, and reproducing the results of data-analysis pipelines. For this purpose, the application code should be subject to revision source control, and the

implementation of pipelines should be encoded formally using a standard representation such as the directed acyclic graphs (DAGs) used by the Condor scheduler.

- *Computational resources* (COMP-6). The computational infrastructure must also ensure the availability of sufficient computational resources for the data-analysis tasks, either in-house or with the data-analysis agents. Undoubtedly, Grid-enabling the LISA computational infrastructure will play a large part in this endeavor.

4.2.2 Intermediate milestones and development tasks

High-level milestones for Computational Infrastructure Total estimated effort for Phase A: 1.15 FTE		Deploy and demonstrate *initial computational infrastructure* testbed for use as testground in data-analysis application development.	Deploy and demonstrate *prototype computational infrastructure* for extensive testing of serial and parallel data-analysis algorithms and pipelines and for accurate characterization of computational requirements. Assess computational requirements for infrastructure.	Deliver specifications for *final production computational* infrastructure. Deploy final computational infrastructure and secure computational resources. Prepare for transition to production infrastructure interfaced with data center.
Intermediate milestones		**Phase A**	**Phase B**	**Phase C**
COMP-1	Data storage and delivery facilities and libraries Estimated effort for Phase A: 0.15 FTE Priority for Phase A: middle	Simplified data storage and delivery facilities and libraries.	Robust and scalable facilities and libraries.	Production facilities and libraries.
COMP-2	Data reduction and conditioning libraries Estimated effort for Phase A: 0.15 FTE Priority for Phase A: middle	Simple libraries founded on high-level model of TDI reduction and on evolving designs for phasemeter, laser, and proof-mass subsystems.	Develop a realistic and increasingly inclusive theoretical model of low-level measurement and data reduction (including ranging, telemetry, housekeeping, instrument vetos, TDI vetos, quality cuts), on the basis of evolving designs for payload and spacecraft subsystems. Deliver data reduction and conditioning libraries based on this model.	Deliver specifications for production libraries. Include insight from testing of realistic models of payload and spacecraft subsystems. Deploy libraries.
COMP-3	Source and noise simulators Estimated effort for Phase A: 0.50 FTE Priority for Phase A: top	Package existing simulators (*LISA Simulator*, *Synthetic LISA*) for use in testbed computational infrastructure, providing extensive documentation and unit-testing frameworks. Add support for prototype data formats, and interface with source-waveform codes.	Implement low-level model of measurement and data reduction. Include "black-box" modules (employing sampled data and measured characteristics) to model noise and response of selected payload and spacecraft subsystems. Increase variety and availability of simulated gravitational-wave signals. Deliver practical interface to the *LISA Standard Model* (AST-RT-1).	Increase fidelity of simulators to use as LISA "stand-in" for simulation demonstration of LISA data-analysis system. Include "black-box" modules for most payload and spacecraft subsystems.

COMP-4	File formats and data structures Estimated effort for Phase A: 0.25 FTE Priority for Phase A: top	Develop flexible and extensible data formats for high level datasets (e.g., TDI observables) and source waveform repositories and generation codes.	Extend data formats to include auxiliary science-keeping channels and other low-level data, including occurrence and characterization of gaps and glitches. Develop expressive representation for probabilistic, interdependent knowledge about detected sources (i.e., data models), especially targeted at global-fit problem.	Deliver specifications for production data-analysis file formats.
COMP-5	Application guidelines Estimated effort for Phase A: 0.10 FTE Priority for Phase A: middle	Develop initial set of guidelines for contributions to computational-infrastructure testbed and proof-of-concept applications.	Investigate use of formal verification techniques for LISA software. Investigate use of formal representations to ensure the traceability of data-analysis pipelines.	Deliver final guidelines for NASA/ESA data-analysis system and for client (e.g., guest investigator) applications.
COMP-6	Computational resources	N/A	Provide framework to assess computational requirements of data-analysis tasks.	Develop tentative plan for in-house and outsourced computational capabilities; and investigate and plan Grid integration accordingly.

4.2.3 Status

4.2.4 Risk Reduction

4.3 MET-SS and MET-GF – Data analysis algorithms and methods

4.3.1 Background

According to the currently envisaged architecture of LISA data analysis, algorithms and methods can be loosely classified into two classes:

- *source-specific detection and parameter estimation methods* (see Secs. 4.3.2a, 4.3.3a, 4.3.4a), which use parametrized source waveforms (or other looser characterizations of prospective signals) to attribute a probability to the presence of an individual source of given parameters in the LISA data stream;
- *global-fit procedures* (see Secs. 4.3.2a, 4.3.3a, 4.3.4a), which call on the source-specific methods as *plugins,* with the purpose of establishing a global interpretation of the LISA data stream as the superposition of a large number of sources. Thus, global-fit procedures attempt to correct for the bias introduced in the detection and parameter characterization of individual sources by the presence of other signal; as a byproduct, global-fit procedures would also produce “cleaned” data streams (where all confidently detected sources have been subtracted out) for deeper searches or for posing constraints on diffuse backgrounds.

Some of the data analysis tasks are both difficult and *unprecedented*, so this plan emphasizes the exploration of data-analysis *fundamentals* (i.e., theories to tell us how and how well a given task can be performed), as well as the *robustness* of proposed methods (which can be tested with high-fidelity simulated data including faithful models of instrument noise and simulated measurement events such as gaps and glitches). Resource assessment (manpower, storage and computing resources, which may dictate a choice between alternative methods) is also considered in this development area.

4.3.2a Intermediate milestones and development tasks (source-specific methods)

High-level milestones for Source-Specific Detection and Parameter Estimation Methods and Tools Total estimated effort for Phase A: 4.5 FTE		Develop and demonstrate (by extensive simulations on realistic data sets) *proof-of-concept* methods for reliable detection and initial parameter estimation for the sources included in the *minimum* science measurement requirements. Develop and demonstrate (by analytical arguments and simulations on representative data sets) *proof-of-concept* methods for reliable detection, parameter estimation, and concurrent fitting for all sources included in the *full* baseline-mission science measurement requirements.	Develop and demonstrate (by extensive simulations on high-fidelity data sets including realistically complex waveforms and source populations of realistic magnitude) robust *prototype* methods for reliable detection, accurate parameter estimation, and concurrent fitting for all sources included in the *full* baseline-mission science measurement requirements. Assess computational requirements for source-specific data-analysis applications and global-fit plugins.	Deliver specifications for detection and parameter-estimation stand-alone applications and global-fit plugins. Deploy production source-specific applications and plugins.
Intermediate milestones		**Phase A**	**Phase B**	**Phase C**
MET-SS-1	Binaries Note: Galactic and Magellanic binaries; includes verification binaries. (SRD 3.1, 3.2) Estimated effort for Phase A: 1.15 FTE Priority for Phase A: top	Demonstrate proof-of-concept methods. Demonstrate production of catalog down to S/N=20. Determine whether five or six working links are needed to determine polarization and inclination independently of orbital modulation (needed to map SR_S1.3 and SR_S5.2 into a measurement requirement).	Robust methods.	Production methods.
MET-SS-2	MBHs (SRD 3.3, 3.4) Note: includes SMBH binaries ($M_1 > 3\ 10^4\ M_{Sun}$, mass ratio > 0.01) and IMBH binaries ($M_1 < 3\ 10^4\ M_{Sun}$) Estimated effort for Phase A: 0.75 FTE Priority for Phase A: top	Demonstrate proof-of-concept methods. Demonstrate proof of concept for rapid-turnaround data-analysis system to determine time and position of merger two months in advance. Explore practical limits on subtraction of strong MBH signals, and estimate waveform accuracy to be requested from numerical relativity as needed for accurate global fits. (Work in the worst-case scenario of high rates for comparable-mass binaries.)	Robust methods. Develop and demonstrate framework to compare observed signals with waveforms computed in numerical relativity. Demonstrate luminosity-distance measurement.	Production methods.

MET-SS-3	EMRIs Note: $M_1 > 3\ 10^4\ M_{Sun}$, mass ratio < 0.01 (SRD 3.5) Estimated effort for Phase A: 1.25 FTE Priority for Phase A: top	Develop and test proof-of-concept methods for semicoherent and time-frequency (incoherent) detection and parameter estimation; attempt estimates of computational requirements. Explore effects of self confusion.	Robust methods. Extend detection methods to systems where the central object is not a standard GR BH. Develop practical schemes to interpolate between computationally expensive numerical waveforms.	Production methods.
MET-SS-4	Backgrounds (SRD 3.2 and 3.6) Estimated effort for Phase A: 0.75 FTE Priority for Phase A: middle	Investigate proof-of-concept methods to detect and estimate distribution of the Galactic-Magellanic background, and to establish upper limits on the isotropic GW background. Investigate algorithms for separating instrument noise from isotropic and anisotropic GW backgrounds (see also MET-SS-5).	Robust methods. Develop and demonstrate capability to produce sky map of the GW background. Develop framework to estimate fallback of global-fit residuals on background estimates.	Production methods.
MET-SS-5	Unmodeled sources (SRD 3.6) Estimated effort for Phase A: 0.6 FTE Priority for Phase A: middle	Develop and test proof-of-concept methods to identify instrument and non-instrument non-Gaussian bursts (especially TDI-based template-less methods such as the zero-signal solution).	Robust methods. Develop framework to derive burst-signal vetos from catalog of instrument noise transfer functions	Production methods.

4.3.2b Intermediate milestones and development tasks (global-fit methods)

Higher-Level milestones for Global-Fit Procedures Total estimated effort for Phase A: 3.2 FTE		Investigate, develop, and demonstrate (by simulations on representative data sets) alternative *candidate* procedures for the *global-fit* detection and parameter estimation of overlapping signals from different source classes.	Select, develop and demonstrate (by extensive simulations on high-fidelity data sets) robust *prototype* global-fit procedures that can be scaled up to source populations of real magnitude. Assess computational requirements for global-fit pipelines (using source-specific information from MET-SS-B).	Review, encourage, and commission community-based research on astrophysical models needed to secure maximum science payoff from LISA. in crucial areas where early lead-in is needed to secure maximum science payoff from LISA.
Intermediate milestones		**Phase A**	**Phase B**	**Phase C**
MET-GF-1	Fundamental theoretical aspects Estimated effort for Phase A: 1.0 FTE Priority for Phase A: top	Extend standard signal-processing theory to provide reliable estimates of capabilities for the resolution and parameter characterization of individual sources in overlapping ensembles. Understand interpretation of multimodal posterior probability distributions. Understand statistical significance of detections and global fits in signal-dominated data stream. Give operational definitions of confusion noise. Explore theoretical limits on subtraction of strong foreground signals. Understand statistics of residual signals after global-fit iterations.	Continue to refine investigations of Phase A. Determine structure of probabilistic source catalog to be released periodically (Bayesian network? Covariance matrix?).	
MET-GF-2	Algorithms Estimated effort for Phase A: 1.0 FTE Priority for Phase A: top	Devise, implement, and test (on small source ensembles) alternative algorithms for global-fit detection and parameter estimation. Compare iterated-detection and concurrent-detection schemes.	Extend development, testing, and selection of global-fit methods to the full spectrum of LISA sources in full-scale populations, as specified by the *LISA Standard Model* (AST-RT-1). Assess computational requirements and revise architecture and algorithm selection according to projected availability of resources.	

MET-GF-3	Interaction with instrument noise Estimated effort for Phase A: 1.0 FTE Priority for Phase A: top	Examine role of instrument noise diagnostics and independent noise estimates as inputs to the global-fit problem. Explore effects of data gaps and glitches, and requirements to be set on them.	Develop framework and tools to include probabilistic noise estimates in candidate global-fit algorithms. Test with realistic simulated noise.	
MET-GF-4	Interaction with diffuse backgrounds Estimated effort for Phase A: 0.2 FTE Priority for Phase A: middle	Investigate and demonstrate mitigation of diffuse background interference on detection of other sources.	Develop robust mechanism for diffuse-background interference mitigation.	

4.3.3a Status (source-specific methods)

4.3.3b Status (global-fit methods)

4.3.4a Risk Reduction (source-specific methods)

4.3.4b Risk Reduction (global-fit methods)

4.4 AST-WF and AST-RT – Astrophysical models

4.4.1 Background

This development area covers work in astrophysical research leading to *source waveforms or signal models*, as needed for detection (see Secs. 4.4.2a, 4.4.3a, 4.4.4a), as well as work centering on *event rates and populations*, as needed to inform measurement requirements and search strategies, and to set up *inverse-problem frameworks* to extract the maximum astrophysical information from the LISA data (see Secs. 4.4.2b, 4.4.3b, 4.4.4b).

4.4.2a Intermediate milestones and development tasks (waveforms and models)

Higher-level milestones for Astrophysical Waveforms and Models Total estimated effort for Phase A: 0.7 FTE		Track progress in community-based research on astrophysical models needed to secure maximum science payoff from LISA.	Review, encourage, and commission community-based research on astrophysical models needed to secure maximum science payoff from LISA. in crucial areas where early lead-in is needed to secure maximum science payoff from LISA.	Compile complete set of astrophysical models to allow the optimal extraction of scientific information from the LISA data.
Intermediate milestones		**Phase A**	**Phase B**	**Phase C**
AST-WF-1	Binaries Estimated effort for Phase A: 0.1 FTE Priority for Phase A: middle	Track progress.	Develop parametrized waveforms for dirty (tidally-dissipating and mass-transferring) binary systems; evaluate effects off dirty waveforms on system detection and parameter estimation.	Extend Phase B work.
AST-WF-2	MBHs Estimated effort for Phase A: 0.1 FTE Priority for Phase A: middle	Track progress.	Model possible EM-counterpart emission from MBH-binary environments; prepare for interaction with non-GW observatories.	Extend Phase B work.
AST-WF-3	EMRIs Estimated effort for Phase A: 0.1 FTE Priority for Phase A: middle	Track progress.	Model possible EM-counterpart emission from MBH-binary environments; prepare for interaction with non-GW observatories.	Extend Phase B work.
AST-WF-4	Cosmic background Estimated effort for Phase A: 0.1 FTE Priority for Phase A: middle	Track progress.	Review possible levels and spectra for cosmic backgrounds. Establish formalism to compare the cosmological GW background to known diffuse electromagnetic (EM) backgrounds (e.g., the CMB) and characterize the isotropy of the diffuse cosmic background.	Extend Phase B work.

AST-WF-5	Burst sources (Include cosmic strings, VMO supernovae, "occasional" Sgr A* stellar passages.) Estimated effort for Phase A: 0.3 FTE Priority for Phase A: middle	Catalog possible galactic and cosmological sources of exotic or speculative nature.	Model sources. Add newly proposed sources.	Extend Phase A and B work.
AST-WF-6	GW propagation	N/A	Model propagation effects on GWs from density inhomogeneities (weak lensing) to verify effects on parameter estimation (esp. use of MBHs as standard candles).	Extend Phase B work.

4.4.2b Intermediate milestones and development tasks (event rates, populations, and inverse problems)

Astrophysical event rates, populations, and inverse problems Total estimated effort for Phase A: 3.7 FTE		Characterize astrophysical event rates and populations as needed to finalize science measurement requirements, to prepare for on-orbit commissioning, and to inform the development of data-analysis techniques.	Increase fidelity of event-rate and population models. Formulate inverse problems to obtain the maximum scientific payoff from the LISA data.	Extend phase B work.
Intermediate milestones		**Phase A**	**Phase B**	**Phase C**
AST-RT-1	Verification binaries Estimated effort for Phase A: 0.3 FTE Priority for Phase A: top	Identify known binaries, measure their properties, characterize expected statistical significance of detection.	Commission new surveys if more binaries are needed for system-verification purposes.	N/A
AST-RT-2	MBHs Estimated effort for Phase A: 1.5 FTE Priority for Phase A: middle/top	Improve estimates of event rates for SMBHs (needed especially for estimating the accuracy needed in foreground subtraction). Improve estimates of parameter distributions (e.g., spins, eccentricities) as input to the study of detection methods.	Extend phase A work.	Extend phase A and B work.
AST-RT-3	EMRIs Estimated effort for Phase A: 1.0 FTE Priority for Phase A: middle/top	Improve lower bound on event rates for EMRIs (needed for finalizing measurement requirements).	Formulate inverse problem to characterize the demographics of dense galactic nuclei and disks from a catalog of detected EMRI events.	Extend phase A and B work.

AST-RT-4	Galactic/Magellanic background Estimated effort for Phase A: 0.6 FTE Priority for Phase A: middle	Develop realistic galactic binary source simulations involving full galactic populations that produce a diffuse galactic binary background at low frequencies.	Formulate inverse problem for Galactic/Magellanic WD populations. Develop new population-synthesis codes or instrument existing codes as input for Standard Model and test for inverse-problem framework.	Extend phase A and B work.
AST-RT-5	Burst sources	N/A	Include cataloged burst sources in the LISA Standard Model.	Extend phase B work.
AST-RT-6	The *LISA Standard Model* Estimated effort for Phase A: 0.3 FTE Priority for Phase A: top	Devise plan to track the current understanding of astrophysical populations and parameter distributions; interface with GW simulators to generate full LISA dataset including all source types.	Begin publishing the LISA standard model to data-analysis agents. Keep updating as astrophysical knowledge progresses. Interface with improved source-simulation codes as they become available.	Keep updating LISA standard model.

4.4.3a Status (waveforms and models)

4.4.3b Status (event rates, populations, and inverse problems)

4.4.4a Risk Reduction (waveforms and models)

4.4.4b Risk Reduction (event rates, populations, and inverse problems)

4.5 GR – General Relativity

4.5.1 Background

This development area focuses on the outstanding problems of general relativity whose timely solutions will enable the efficient and accurate detection and parameter characterization of LISA sources, as well as the extraction of the maximum scientific information from the observations. These problems include:

- The generation of reliable calculations of the great variety of gravitational waveforms emitted in the LISA frequency band by quasi-equal mass compact binary systems. This will enable identification and astrophysical interpretation of such signals in the LISA data streams, and typing and mapping the distributions of these objects throughout the galaxy;
- The development of techniques for mapping the strongly warped spacetime in the vicinity of a spinning black hole, using LISA observations of the inspirals of extreme mass ratio binaries (EMRIs). EMRIs involve solar-mass compact objects such as white dwarfs or black holes, coalescing with supermassive black holes. The direct observation of this space-time geometry will be an unprecedented experimental verification of Einstein's theory of general relativity and its relativistic field theory of gravitation.
- The quantification of possible fundamental tests of gravitational theory that can be made with LISA, confronting, in the strong gravity regimes of astrophysics, the predictions of general relativity with the predictions of a variety of different modern gravitation theories (*e.g.*, parameterized post-Newtonian theory, the classical limit of string theories, or expected cosmological backgrounds arising from different quantum big bang models).

4.5.2 Intermediate milestones and development tasks (event rates, populations, and inverse problems)

General-relativistic waveforms and models Estimated effort for Phase A: 2.75 FTE		Investigate general-relativistic waveforms and models, as needed to finalize measurement requirements, to inform the development of data-analysis techniques, and to indicate need for community-based effort where major advances are needed.	Develop and deliver tools to efficiently compute *realistic* general-relativistic waveforms and models, as needed to develop and test source-specific and global-fit data-analysis applications.	Develop and deliver tools to efficiently compute *high-fidelity* general-relativistic waveforms and models, as needed for production data-analysis system and to secure maximum science payoff (optimal detection rates, accurate parameter studies).
Intermediate milestones		**Phase A**	**Phase B**	**Phase C**
GR-1	Binaries	N/A	Galactic binary waveforms that include GW-light speed difference, scalar-tensor theories, etc.	Extend phase B work.
GR-2	MBHs Estimated effort for Phase A: 1.0 FTE Priority for Phase A: top	Assess quality of gravitational waveforms produced from the post-Newtonian approximation for binary inspirals, and identify further need of technical development in this area. Clarify instrument requirements for detection latency, protected periods, low frequency sensitivity.	Generate reliable calculations of the great variety of waveforms emitted by comparable-mass systems, possibly including effects of eccentricity, spins, and higher-order post-Newtonian corrections.	Secure and characterize accurate numerical relativity calculations of binary mergers. Develop inverse techniques.

GR-3	EMRIs Estimated effort for Phase A: 1.0 FTE Priority for Phase A: middle/top	Develop a formalism to include (and parametrize) non-central-black-hole and non-general-relativistic effects into the waveforms.	Develop practical and accurate methods to compute radiation-reaction forces and the resulting orbital evolutions of EMRI systems. Develop practical methods to compute the corresponding waveforms as quickly as possible. Provide a rigorous formulation of no-hair tests with EMRIs (non-Kerr or non-GR central objects). Develop a formalism to deduce tidal coupling of the compact object to the central object.	Develop perturbative radiation reaction calculations to the point of producing accurate EMRI waveforms.
GR-4	Backgrounds	N/A	N/A	N/A
GR-5	Burst sources	N/A	N/A	N/A
GR-6	Fundamental tests Estimated effort for Phase A: 0.75 FTE Priority for Phase A: middle/top	Quantify other possible fundamental tests confronting the predictions of GR with those of alternative theories.	Develop tests.	Extend phase B work.

5. Management Plan

6. References

(…to be completed…)

Barack, L., and Cutler , C. (2004a), "LISA Capture Sources: Approximate Waveforms, Signal-to-Noise Ratios, and Parameter Estimation Accuracy," Phys. Rev. D **69**, 082005.

Barack, L., and Cutler , C. (2004b), "Confusion Noise from LISA Capture Sources," Phys. Rev. D **70**, 122002.

Farmer, A. J., and Phinney, E. S. (2003), "The Gravitational Wave Background from Cosmological Compact Binaries," Mon. Not. Roy. Astron. Soc. **346**, 1197.

Gair, J. R., et al. (2004), "Event rate estimates for LISA extreme mass ratio capture sources," Class. Quant. Grav. **21**, S1595.

Kellermann, K., and Sheats, B., eds. (1983), "Serendipitous Discoveries in Radio Astronomy", Proceedings of the Workshop held at the National Radio Astronomy Observatory, Green Bank, WV, May 4-6, 1983.

Phinney, E. S., for LISA Working Group 1 (2002), "LISA Science Requirements", presentation given on March 7, 2002.

Shaddock, D. A., et al. (2004), "Post-processed time-delay interferometry for LISA," Phys. Rev. D **70**, 081101.

Thorne, K. S. (1987) "Gravitational Radiation" in *Three Hundred Years of Gravitation* (S. Hawking and W. Israel, eds.), pp. 330–458, Cambridge University Press, Cambridge.