Document for Challenge 1
Draft v1.0

The taskforce
http://www.tapir.caltech.edu/dokuwiki/listwg1b:home

July 20, 2006

Chapter 1

Introduction

At the LISA International Science Team (LIST) meeting of December 2005, in Pasadena, the Working Group on Data Analysis (LIST-WG1B) decided to embark in the organisation of several rounds of mock data challenges (MLDC), with the dual purpose of:

1. fostering the development of LISA data analysis tools and capabilities, and
2. demonstrating the technical readiness already achieved by the gravitational-wave community in distilling a rich science payoff from the LISA data output.

The LISA Mock Data Challenges were proposed and discussed at meetings organized by the US and European LISA Project that were attended by a broad cross section of the international gravitational-wave community. These challenges are meant to be blind tests, but not really a contest.

This (working) document contains a summary of the key activities of the Mock LISA Data Challenge (MLDC) Taskforce that has been has been charged to formulate challenge problems of maximum efficacy, to establish criteria for the evaluation of the analyses, to develop standard models of the LISA mission (orbit, noises) and of the LISA sources (waveforms, parametrisation), to provide computing tools such as LISA response simulators, source waveform generators, and a Mock Data Challenge file format, and more generally to provide any technical support necessary to the challengers, including moderated discussion forums and a software repository.

The first set of challenge data-sets will be released during the 6th LISA Symposium (June 19-23, 2006, at Goddard Space Flight Center, Greenbelt, Maryland). The challenges will involve the distribution of several taste's, encoded in a simple standard format, and containing combinations of realistic simulated LISA noise with the signals from one or more LISA gravitational-wave sources of parameters unknown to the challenge participants.

It is envisaged that the results of the first MLDCs will be presented to the broad community and discussed in a dedicated session at the 11th Gravitational-Wave Data Analysis Workshop (December 18-21, 2006, at the Albert Einstein Institute, in Golm). The second and third sets of challenge data sets, embodying more ambitious data-analysis problems, will be released in December 2006, with target time-frames for the completion of the analyses in June and December 2007.

1.1 Overall scheme of the challenges

The MLDCs consist in extracting the maximum amount of information about the source(s) that generate gravitational wave signal(s) contained in the (mock) data sets that are distributed. Each round of MLDCs consists of multiple data sets containing signals of different nature and strength.

The overall structure of the challenges is as follows:

Challenge 1 – The goal of this challenge is to foster the development and validation of building blocks of and basic tools for LISA data analysis and tackle analysis of data sets containing a single signal or non-overlapping multiple signals (with one exception, see Chapter 5) embedded in Gaussian and stationary noise with no contribution from galactic and/or extragalactic foregrounds. Training data sets (where all the source parameters are public) will also be provided.

- **Sources**: galactic binaries, verification binaries and massive black hole binaries (only in-spiral portion of the whole coalescence).
- **Release date**: 30 June 2006
- **Due date**: 1 December 2006
- **Details**: provided in Chapter 5
- **Note**: *Data sets containing one EMRI will also be distributed for this first round; however, due to the difficulty of the problem results are requested by the due date of challenge 2, tentatively June 2007*

Challenge 2 – The goal of this challenge is to tackle a global analysis problem within a restricted parameter space in the presence of a galactic foregrounds. As preliminary guideline, the data set shall contain verification binaries, ≈ 100 unknown galactic binaries, a small band with strongly overlapping sources and a few in-spirals from massive black hole binary systems. Challenge data sets containing a single source will also be distributed, for EMRIs, stochastic signals and broad-band short-lived bursts. Training data sets (where all the source parameters are public) will also be provided.

- **Sources**: galactic binaries, verification binaries, massive black hole binary systems, EMRIs, stochastic signals, bursts
- **Release date**: December 2006
- **Due date**: June 2007

Challenge 3 – The goal of this challenge is to test a more realistic data analysis scenario, with multiple sources with more general waveforms and/or more realistic noise contributions (non-stationary, non-Gaussian). Gaps might also be introduced in the data sets. Details are TBD. Training data sets (where all the source parameters are public) will also be provided.

- **Sources**: galactic binaries, verification binaries, massive black hole binary systems, EMRIs, stochastic signals, bursts
- **Release date**: June 2007
- **Due date**: December 2007

Challenge 4 and beyond – TDB

- **Sources**: TBD
- **Release date**:
- **Due date**:

1.2 Challenge organisation and schedule

The challenge data sets with all the relevant information (including the present document), relevant software, and links to other resources are available at `http://astrogravs.nasa.gov.edu/docs/mldc/`.

A member of the Task Force that will not participate in the first round of challenges (John Baker) has produced the data sets and is the repository of the key to disclose the source parameters. They will be made public just after 1 December 2006, the due date of Challenge 1.

After the release of the data sets, the Task Force plans to organise a teleconference with all the participating groups, with date still TBD around the beginning of October, 2006 in order to track progress.

The Challenge-1 due date is 1 December 2006. Results and relevant documents should be submitted at the MLDC website `http://astrogravs.nasa.gov.edu/docs/mldc/` following the link provided.

The Task Force plans to host a teleconference in the time frame 5-8 December (to be confirmed) to discuss results with the participating groups. A face-to-face meeting is also tentatively scheduled to take place at GWDAW which will be held at the Albert-Einstein-Institut (Golm, Germany) 18-21 December, 2006. A half a day session for presentations of the MLDC results from the participating groups and the Task Force is also scheduled for the workshop. Attendance to GWDAW is not a requirement to take part to the MLDC.

1.2.1 Participating to Challenge-1

Participating groups are asked to subscribe to the MLDC Challenge-1 mailing list at:

`https://lists.sourceforge.net/lists/listinfo/lisatools-challenge`

The Task Force encourages participating groups to circulate a brief note with the plans for this round of challenges, in particular which sub-challenges (see Chapter 5 for more details) are going to be tackled and what kind of analysis techniques will be used. These plans will be made public and posted on the MLDC website in order to make all the groups fully aware of the on-going activities.

1.2.2 Reporting results

Participating groups are requested to report the results in the following form:

- Files containing information about the parameters that characterise the signal(s) in the mock data stream in the form of posterior probability density functions, confidence intervals, or any statistical quantity that is relevant for the adopted analysis strategy;
- A technical note describing:
  - The analysis method
  - The pipeline, software implementation, platform(s) on which the analysis was carried out and an estimate of the run times

The participating groups are strongly encouraged (if at all possible) to tag the codes used for "production analysis" and maintain them under version control It is at the participant discretion to make the software available; although the Task Force encourages this approach, this is not a requirement to participate to the MLDCs.

1.2.3 Evaluating results

The Task Force shall formulate a matrix to be used to compare the results of the challenges and address the readiness in tackling the given data analysis task.

1.2.4 Dissemination of the results

The dissemination of the results will include:

- A technical note by each participating group reporting the analysis method and the results of the analysis
- A technical note issued jointly by the Task Force and the participating groups providing a summary of the results and highlighting outstanding issues and
- Papers to be published in the proceedings of GWDAW reporting the results of the analysis by each participating group
- Possibly a joint paper by the Task Force and the participating groups to be published in an international peer review journal (this is most likely appropriate from Challenge-2 onwards)

The Task Force encourages the publication of "method papers" by groups taking part to the MLDCs.

Chapter 2

LISA conventions

2.1 Introduction

The goal of this section is to define a basic set of conventions for describing virtual LISA observatories. The authors of the *LISA Simulator* and *Synthetic LISA* simulation software have committed to adopting these conventions. Early papers on LISA, and even early versions of the simulators, have at times used other conventions. An appendix to this document will contain a Rosetta Wheel that explains how results from other sets of conventions can be mapped onto the new standard.

Where possible, conventions have been adopted that are already in common use. In some cases, such as with the initial reference orbits and TDI variables, the choices incorporate certain approximations. At a later date, when LISA data analysis has reached a greater degree of maturity, it will be necessary to transition to a more refined orbital model, and more complex variants of the basic TDI variables. To paraphrase Einstein, the goal is to make the conventions as simple as possible, but no simpler.

For the impatient reader, the initial pseudo-LISA conventions are as follows: Positions are quoted in Heliocentric Ecliptic Coordinates. Times are those measured by a clock at the Solar barycenter. The LISA constellation is described as a rigidly rotating equilateral triangle with a guiding center that follows a circular orbit in the ecliptic with radius 1 AU. The basic science data are the generation 1.5 TDI variables that account for the rotation of the constellation.

2.2 Coordinate System

We imagine that ultimately LISA's motion will be most conveniently expressed using ecliptic coordinates centered at the Solar System Barycenter. That is because the JPL ephemeris will be required to calculate the influence of the Earth, Moon, Jupiter, etc. on the precise motions of the LISA satellites, and the JPL ephemeris works with SSB-centered coordinates. However the LISA model currently used by the *LISA Simulator* and *Synthetic LISA* do not include such planetary influences on the spacecraft, and if one idealizes away planetary influences then it is equivalent to working in Heliocentric Ecliptic Coordinates (HEC). Therefore for simplicity we will use HEC coordinates, and for the purposes of this document, we can simply imagine the Sun fixed at the origin of the HEC coordinate system, with the z axis pointing toward the north ecliptic pole; the $x-y$ plane lies in the ecliptic and the x axis points towards the first point of Aries (where the Earth, as seen from the Sun, will be located on September 22nd); the y axis completes a right-handed orthogonal coordinate system. Sky positions in the HEC system are quoted as ecliptic longitude and ecliptic latitude. When referring to the location of objects far from the solar system (*i.e.*

typical LISA sources) the HEC system coincides with the Geocentric Ecliptic Coordinate (GEC) system. Sky positions in the GEC system are quoted as celestial latitude and celestial longitude. The GEC system is in turn related to the standard Geocentric Equatorial Coordinates (GC) by a simple coordinate rotation about the x-axis (which is shared by all three coordinate systems). Sky positions in the GC system are quoted as Right Ascension (α) and Declination (δ). To be concrete, we use J2000 coordinates, in which the obliquity of the ecliptic (*i.e.* the angle between the ecliptic plane and the equatorial plane) is equal to $\epsilon = 23.439291^\circ$. Hopefully LISA will launch before we need to move to J2050 coordinates.

2.3 LISA orbits

The orbits of the pseudo-LISA spacecraft are defined in conformity with the Appendix of Ref. (2) (and in conformity with the *LISA Simulator*). Though not stated explicitly, Ref. (2) employs the HEC system of coordinates. The initial reference orbit is defined by truncating the exact Keplerian orbits at first order in the eccentricity e. The coordinates of each spacecraft are then given by the expressions

$$\begin{aligned} x &= a\cos(\alpha) + a\,e\left(\sin\alpha\cos\alpha\sin\beta - (1+\sin^2\alpha)\cos\beta\right), \\ y &= a\sin(\alpha) + a\,e\left(\sin\alpha\cos\alpha\cos\beta - (1+\cos^2\alpha)\sin\beta\right), \\ z &= -\sqrt{3}\,a\,e\cos(\alpha-\beta), \end{aligned} \tag{2.1}$$

where $\beta = 2(n-1)\pi/3 + \lambda$ ($n = 1, 2, 3$) is the relative orbital phase of each spacecraft in the constellation, a is the semi-major axis of the guiding center, and $\alpha(t) = 2\pi f_m t + \kappa$ is the orbital phase of the guiding center. At this order of approximation the spacecraft form a rigid equilateral triangle with side length $L = 2\sqrt{3}ae$. Setting $e = 0.00965$ and $a = 1$ AU yields the standard $L = 5 \times 10^6$ km arm lengths.

Notice that by keeping only linear terms in the eccentricity we are neglecting the variation in the optical path length that would be present if the full Keplerian orbits were used.[1] The reason for this truncation is twofold. First, it makes very little difference to the instrument response, and second, there are periodic and secular effects on the orbits from other solar system bodies (notably Earth and Jupiter) that are comparable in size to the higher order Keplerian corrections. The precise form of the orbital perturbations will depend on when LISA is launched and the final orbital injection, so it is difficult to define a convention that is meaningful beyond leading order in e. At a later date the reference orbit could be updated to a full ephemeris model based on a particular launch date and orbital injection.

The parameters κ and λ set the initial location and orientation of the LISA constellation. They are related to the parameters $\bar{\phi}_0$ and α_0 used by Cutler (32) according to the mapping

$$\begin{aligned} \bar{\phi}_0 &= \kappa \\ \alpha_0 &= \frac{3\pi}{4} + \kappa - \lambda, \end{aligned} \tag{2.2}$$

and to the parameters η_0 and ξ_0 used by *Synthetic LISA* (4; 7) by the mapping

$$\begin{aligned} \eta_0 &= \kappa, \\ \xi_0 &= 3\pi/2 - \kappa + \lambda, \\ sw &< 0. \end{aligned}$$

[1] However: the LISA Simulator and Synthetic LISA actually use expressions accurate to order e^2 for the positions. Synthetic LISA uses approximate arm lengths accurate to order e plus the effects of rotation (i.e., pointing ahead), while the LISA Simulator uses arm lengths accurate to order e^2 plus the effects of pointing ahead.

(Setting *sw* to a negative value has the effect of exchanging spacecraft 2 and 3, which would be otherwise reversed with respect to the LISA Simulator.) The default choice is to set $\kappa = \lambda = 0$ at barycentric time $t = 0$. This choice places LISA at the first point of Aries, with spacecraft 1 directly below the guiding center in the southern ecliptic hemisphere.

[[**We should also give the conversion that maps to the Pre-Phase A report and other earlier works such as Peterseim, Jennrich, and Danzmann, CQG 13, 279 (1996); Schilling, CQG 14, 1513 (1997); Peterseim, Jennrich, and Danzmann, CQG 14, 1507 (1997); plus others from the 1997 CQG proceedings.**]]

2.4 LISA responses

The basic LISA response to gravitational waves is taken to be the *phase response* Φ_{ij} used in the *LISA Simulator* and discussed in Sec. II of Ref. (2) [see especially Eqs. (4)–(13) and (22)], or equivalently the *fractional frequency response* $y_{slr}^{\rm gw}$ used in *Synthetic LISA* and discussed in Sec. II B of Ref. (7). Here i and s identify the transmitting spacecraft, j and r the receiving spacecraft for each phase measurement, l is a redundant link index given by $l = 1 : s = 3 \to r = 2$; $2 : 1 \to 3$; $3 : 2 \to 1$; -1 (or $1'$): $2 \to 3$; -2 (or $2'$): $3 \to 1$; and -3 (or $3'$): $1 \to 2$.

The phase and fractional frequency formalisms are equivalent, and related by a simple time integration. *Both data streams will be distributed for the challenges.* The frequency measurements have the advantage of being directly proportional to the gravitational strain; the phase measurements have the advantage of representing more closely the actual output of the LISA phasemeters.

2.5 TDI observables

At present it appears that Time Delay Interferometry (8) will be needed to cancel laser phase noise (arm locking may soften the requirements, but is unlikely to dispense with the need for TDI). *We will adopt the modified Time Delay Interferometry variables (TDI 1.5) (10; 9), as defined below, as the standard pseudo-LISA data outputs.* The modified TDI variables are a nice compromise between the unrealistically simple Michelson variables that are swamped by laser phase noise, and the complicated second generation TDI variables that are designed to cancel laser phase noise in an array that both rotates and flexes. The modified TDI variables fit nicely with the order e truncation of the spacecraft orbits, as the TDI 1.5 scheme is able to account for rotation but not flexing (9).

We define the standard TDI observables following the *Synthetic LISA* (4; 7) naming scheme and sign conventions (see also the *Synthetic LISA* file `lisasim-tdi.cpp`). All of these can be used both as frequency and phase observables by replacing y_{slr} measurements with Φ_{ij} measurements. See the TDI Rosetta Stone (6) for translations between index notations (in particular, the primed indices of Ref. (10) correspond to positive indices in the *Synthetic LISA* usage).

- First-generation TDI (TDI 1.0): the *Sagnac* observables α, β, γ ("centered," respectively, on spacecraft 1, 2, 3, as all following sets of three), and the *symmetrized Sagnac* observable ζ, as defined in Ref. (8). No need to define the eight-pulse observables (Michelson, etc.), which are the same as in modified TDI.

- Modified TDI (TDI 1.5): the *unequal-arm Michelson* observables X, Y, Z; the *relay* observables U, V, W; the *monitor* observables E, F, G; the *beacon* observables P, Q, R; the *Sagnac* observables α_1, α_2, α_3; and the *symmetrized Sagnac* observables ζ_1, ζ_2, ζ_3 as defined in Ref. (10).

- Second-generation TDI (TDI 2.0): the *unequal-arm Michelson* observables X_1, X_2, X_3; the *relay* observables U_1, U_2, U_3; the *monitor* observables E_1, E_2, E_3; the *beacon* observables P_1, P_2, P_3 as defined in Ref. (10).
- Optimal TDI observables: in first-generation TDI, A, E, and T as defined in terms of α, β, γ in Ref. (11); in modified TDI, $\bar{A}$, $\bar{E}$, $\bar{T}$ as defined in terms of α_1, α_2, α_3 in Ref. (12).

Note also that there is a naming conflict here between the first-generation spacecraft-1–centered monitor observable and the first-generation spacecraft-2–centered optimal observable.

TDI Data Streams		
variable	descriptor (phase)	descriptor (fractional frequency offset)
TDI-1.0 Sagnac observables		
α, β, γ	`alphap`, `betap`, `gammap`	`alphaf`, `betaf`, `gammaf`
ζ	`zetap`	`zetaf`
TDI-1.0 *optimal* observables ("`O`" for "optimal")		
A, E, T	`AOp`, `EOp`, `TOp`	`AOf`, `EOf`, `TOf`
TDI-1.0 and TDI-1.5 unequal-arm–Michelson observables		
X, Y, Z,	`Xp`, `Yp`, `Zp`	`Xf`, `Yf`, `Zf`
TDI-1.0 and TDI-1.5 *relay*, *monitor*, and *beacon* observables		
U, V, W	`Up`, `Vp`, `Wp`	`Uf`, `Vf`, `Wf`
E, F, G	`Ep`, `Fp`, `Gp`	`Ef`, `Ff`, `Gf`
P, Q, R	`Pp`, `Qp`, `Rp`	`Pf`, `Qf`, `Rf`
TDI-1.5 Sagnac observables		
α_1, α_2, α_3	`alpha1p`, `alpha2p`, `alpha3p`	`alpha1f`, `alpha2f`, `alpha3f`
ζ_1, ζ_2, ζ_3	`zeta1p`, `zeta2p`, `zeta3p`	`zeta1f`, `zeta2f`, `zeta3f`
TDI-1.5 *optimal* observables ("`B`" for "bar")		
$\bar{A}$, $\bar{E}$, $\bar{T}$	`ABp`, `EBp`, `TBp`	`ABf`, `EBf`, `TBf`
TDI-2.0 unequal-arm–Michelson observables		
X_1, X_2, X_3,	`X1p`, `X2p`, `X3p`	`X1f`, `X2f`, `X3f`
TDI-2.0 *relay*, *monitor*, and *beacon* observables		
U_1, U_2, U_3	`U1p`, `U2p`, `U3p`	`U1f`, `U2f`, `U3f`
E_1, E_2, E_3	`E1p`, `E2p`, `E3p`	`E1f`, `E2f`, `E3f`
P_1, P_2, P_3	`P1p`, `P2p`, `P3p`	`P1f`, `P2f`, `P3f`

2.5.1 Data Streams for Challenge 1

For Challenge 1, the Challenge data sets contain the modified TDI observables $X(t)$, $Y(t)$, and $Z(t)$. Two versions of each data set will be provided: the phase response version produced by the *LISA Simulator* and the frequency response version produced by *Synthetic LISA*.

2.6 Instrumental Noise

The model of the instrumental noise that is adopted for Challenge 1 is as follows:

- No laser phase noise (we assume perfect cancellation with TDI)
- White phase optical noise, with one-sided spctral density given by

$$S_{n,o}^{1/2}(f) = 20 \times 10^{-12} \mathrm{m\,Hz}^{-1/2}\,; \tag{2.3}$$

- White+red acceleration noise, with with one-sided spctral density given by

$$S_{n,a}^{1/2}(f) = 2 \times 10^{-15} \left[1 + \left(\frac{10^{-4}\,\mathrm{Hz}}{f} \right)^{-2} \right]^{1/2} \mathrm{m\,s^{-2}\,Hz^{-1/2}}\,. \tag{2.4}$$

Chapter 3

Sources and waveform conventions

3.1 Introduction

This chapter contains a description of the conventions used to describe the two gravitational wave polarisations and the model we adopted for the different source classes whose signals are present in the mock data sets.

3.2 LISA sources

Each class of gravitational wave sources can be described by several equivalent sets of parameters, but no one set of parameters serves as a natural basis for describing all gravitational wave sources. For example, a typical galactic binary system is described by seven parameters, which are conventionally taken to comprise the amplitude, frequency, HEC colatitude, HEC longitude, inclination, polarization angle and initial orbital phase. Some galactic binaries may require additional parameters such as eccentricity and first and second derivatives of the frequency. Other systems, such as Extreme Mass Ratio Inspirals (EMRIs) are described by a much larger set of parameters since these systems have fewer (approximate) constants of motion. In particular, spin-orbit coupling and perihelion precession render familiar parameters such as inclination and polarization angle meaningless. On the other hand, while the set of parameters that describe an EMRI can be used to describe a galactic binary (since, e.g., the stellar-mass objects will have *some* spin), there is no way to recover all of these parameters from the observed galactic binary signal. In terms of a Fisher information matrix description, the 17 parameter set of EMRI parameters would yield 10 zero eigenvalues when applied to a monochromatic galactic binary. Thus, it is natural to imagine a scheme whereby different source classes are described by different numbers of parameters and different parameter bases. The division into source classes is subtle, and will ultimately be something that will have to be incorporated into the data analysis algorithms themselves. A good example is a slowly evolving galactic binary: with just one year of data it may be well described by 7 parameters, but with three years of data it may be necessary to employ an 8 parameter model that includes the frequency derivative.

For the initial Challenges, all the sources we consider are binaries, but with sufficiently different parameter values and characteristics that further classify them into a few subcategories: (1) Non-evolving, circular stellar-mass (SM) binaries; (2) Slowly evolving circular SM binaries; (3) Non-evolving elliptical SM binaries; (4) Non-spinning, circular-orbit, massive black hole binaries; (5) Spinning, circular-orbit, massive black hole binaries undergoing simple precession; (6) EMRIs. There are other sub-categories that could be added to this list, but for the first Challenges we will

start with this restricted set of possibilities. Likewise, in addition to binary systems, one should be defining conventions for cosmic string bursts, primordial backgrounds, etc, but this is left for later work.

Subsections 3.3, 3.4, and 3.5 below deal with stellar-mass binaries, SMBH inspirals, and EMRIs, respectively. To make the current version of this document, separate manuscripts by different authors were basically concatenated to make these three subsections. While some small effort has been made to achieve a consistency of presentation among these subsections, we warn the reader in advance that *variable names and notations are not consistent between subsections 3.3, 3.4, and 3.5*, though hopefully they are at least consistent *within* each of those subsections. For example, in 3.4 "λ" stands for the ecliptic longitude of the source, while in 3.5 "λ" stands for angle between the SMBH spin vector $\vec{S}$ and the orbital angular momentum $\vec{L}$.

3.2.1 Conventions for Units, Sky Location, and Polarization

Throughout this document we use geometrical units where $G = c = 1$.

We follow Ref. (2) (and the *LISA Simulator*) in describing the sky location of gravitational-wave sources by the unit vector $\hat{n}$,

$$\hat{n} = \sin\theta\cos\phi\,\hat{x} + \sin\theta\sin\phi\,\hat{y} + \cos\theta\,\hat{z}\,, \tag{3.1}$$

(where θ and ϕ are the J2000 *ecliptic colatitude* and *longitude*, the latter measured from the vernal point, aligned with the $\hat{x}$ axis in our convention). The corresponding gravitational radiation is modeled as a plane wave in a transverse-traceless gauge, propagating in the $-\hat{n}$ direction in the HEC frame. The surfaces of constant phase are then given by $\xi = t + \hat{n}\cdot\mathbf{x} = \text{const}$. A generic gravitational wave can be decomposed into two standard polarization states,

$$\mathbf{h}(\xi,\hat{n}) = h_+(\xi)\,\mathbf{e}^+(\hat{u},\hat{v}) + h_\times(\xi)\,\mathbf{e}^\times(\hat{u},\hat{v})\,, \tag{3.2}$$

where $\mathbf{e}^+$ and $\mathbf{e}^\times$ are the polarization tensors

$$\begin{aligned} \mathbf{e}^+ &= \hat{u}\otimes\hat{u} - \hat{v}\otimes\hat{v}, \\ \mathbf{e}^\times &= \hat{u}\otimes\hat{v} + \hat{v}\otimes\hat{u}, \end{aligned} \tag{3.3}$$

and where

$$\begin{aligned} \hat{u} &= \cos\theta\cos\phi\,\hat{x} + \cos\theta\sin\phi\,\hat{y} - \sin\theta\,\hat{z}, \\ \hat{v} &= \sin\phi\,\hat{x} - \cos\phi\,\hat{y}\,. \end{aligned} \tag{3.4}$$

The angles θ and ϕ coincide with the θ_s and ϕ_s used by Cutler (32), and are related to the β (J2000 *ecliptic latitude*) and λ (J2000 ecliptic longitude from the vernal point) used in *Synthetic LISA* (4; 7) by $\theta = \frac{\pi}{2} - \beta$, $\phi = \lambda$.

The remaining pieces of information required to specify the geometry of a binary system depends on whether orbital precession is significant or not. For classes (1) through (4) in the classification scheme given above, orbital precession is not significant.

3.2.2 Basic conventions for binaries

Again, all the sources considered in the initial Challenges are binaries (stellar-mass binaries, MBH binaries, and EMRIs). Our convention is that M_1 is the larger mass and M_2 is the smaller one. From these two quantities one forms various other useful combinations: the total mass is $M \equiv M_1 + M_2$,

the mass difference is $\delta m = m_1 - m_2$, the reduced mass is $\mu \equiv M_1 M_2/M$, the symmetric mass ratio is $\eta \equiv \mu/M$, and the chirp mass is $\mathcal{M} \equiv \mu^{3/5} m^{2/5} = M_1 M_2 (M_1 + M_2)^{-1/3}$.

Note for simplicity in this document (and in the initial Challenges) we are treating the background spacetime as Minkowski space, not Robertson-Walker. For sources at cosmological distances, one should interpret all masses in this document as "redshifted masses;" e.g., M really stands for $M_z = M(1+z)$, where M is the locally measured mass and $1+z$ is the redshift factor, and similarly for M_1,M_2,μ and $\mathcal{M}$. Likewise, the distance to the source D should be interpreted as the "luminosity distance" D_L (36).

3.2.3 Geometry for Non-precessing Binaries

Non-precessing sources can have their orbital orientation described in terms of inclination to the line of sight and polarization angle. For circular orbits the later quantity is equivalent to the longitude of the ascending node. For elliptical orbits the polarization angle is related to a combination of the argument of pericenter and the longitude of the ascending node; the precise mapping was derived by Wahlquist, and will be included in a later draft.

For circular orbits, the polarization angle is defined relative to the *principal polarization axes* $\hat{p}$ and $\hat{q}$ of the source,

$$\mathbf{h}(\xi, \hat{n}) = h_+^S(\xi)\, \boldsymbol{\epsilon}^+(\hat{p}, \hat{q}) + h_\times^S(\xi)\, \boldsymbol{\epsilon}^\times(\hat{p}, \hat{q}), \tag{3.5}$$

with

$$\begin{aligned} \boldsymbol{\epsilon}^+ &= \hat{p} \otimes \hat{p} - \hat{q} \otimes \hat{q}, \\ \boldsymbol{\epsilon}^\times &= \hat{p} \otimes \hat{q} + \hat{q} \otimes \hat{p}\,. \end{aligned} \tag{3.6}$$

we can go back to the general decomposition (3.2) by setting

$$h_+(\xi) = \cos(2\psi)\, h_+^S(\xi) + \sin(2\psi)\, h_\times^S(\xi), \tag{3.7}$$

$$h_\times(\xi) = \cos(2\psi)\, h_\times^S(\xi) - \sin(2\psi)\, h_+^S(\xi), \tag{3.8}$$

where $\psi = -\arctan(\hat{v}\cdot\mathbf{p}/\hat{u}\cdot\mathbf{p})$ is the *source polarization angle*. For a binary system, the inclination angle ι is defined as the angle between the line of sight $\hat{n}$ and the orbital angular momentum vector of the binary $\mathbf{L}$, so that $\iota = \arccos(\hat{L}\cdot\hat{n})$. These variables are related to the (θ_L, ϕ_L) used by Cutler (32) by

$$\begin{aligned} \iota &= \arccos\left(\cos\theta_L \cos\theta_s + \sin\theta_L \sin\theta_s \cos(\phi_s - \phi_L)\right)\,, \\ \psi &= \arctan\left(\frac{\cos\theta_s \sin\theta_L \cos(\phi_s - \phi_L) - \cos\theta_L \sin\theta_s}{\sin\theta_L \sin(\phi_s - \phi_L)}\right)\,, \end{aligned} \tag{3.9}$$

and to the (ψ, ι) used in *Synthetic LISA* (4; 7) by $\iota = \iota_{(\mathrm{SL})}$, $\psi = -\psi_{(\mathrm{SL})}$.

LISA Simulator	Synthetic LISA	Cutler 1998	descriptor	units (first is standard)
sky position				
θ	$\pi/2 - \beta$	θ_s	`EclipticColatitude`[1]	radians, degrees
$\pi/2 - \theta$	β	$\pi/2 - \theta_s$	`EclipticLatitude`[1]	radians, degrees
ϕ	λ	ϕ_s	`EclipticLongitude`	radians, degrees
non-precessing binary geometry				
ι	ι	see Eq. (3.9)	`BinaryInclination`	radians, degrees
ψ	$-\psi$	see Eq. (3.9)	`BinaryPolarization`	radians, degrees

[1]Only one out of `EclipticColatitude` and `EclipticLatitude` required.

3.2.4 Geometry for Precessing Binaries

Not needed yet.

3.3 Galactic binaries

Galactic binaries will ultimately be provided as two distinct types of sources in the Mock LISA Data Challenge—a bulk signal from the Galaxy, and resolvable binaries. In the first instance the source descriptors will be applicable to the model Galaxy as a whole, while in the second instance the source descriptors will be physical parameters of the individual binaries. For Challenge-1, we consider only the latter case: searches for individually resolvable binaries. However even in this case it may be useful to consider the distribution from which the individual binaries are chosen. The distribution we used is described in the next subsection.

3.3.1 Galactic model

In constructing Challenge data sets, Galactic binary parameters are chosen at random based on our model Galactic population density (and subject to constraints on signal strength and frequency range described below).

The bulk properties of the Galaxy can be broken down into morphology and composition. The morphology of a Galaxy is best described in terms of a population density function. Two population densities are currently considered in the literature. The double exponential used by Hils, Bender & Webbink (1990) and Timpano, Rubbo & Cornish (2005) is described entirely by central density, ρ_0, radial scale, R_0, and scale height, z_0, and has the form (in Galactocentric cylindrical coordinates):

$$\rho(\mathbf{r}) = \rho_0 e^{-R/R_0} e^{-z/z_0}. \tag{3.10}$$

The "cuspy" exponential-sech^2 used by Nelemans *et al.* (2001), Benacquista, DeGoes & Lunder (2004), and Edlund *et al.* (2005) includes the possibility for a central cusp and so has an additional parameter c:

$$\rho(\mathbf{r}) = \rho_0 r^{-c} e^{-R/R_0} \mathrm{sech}^2(-z/z_0). \tag{3.11}$$

In the current usage, $c = 0,\ 1$, although any real number within this range is acceptable. The total number of binaries N is related to ρ_0 and can be found by integrating the density over all space.

The composition of the Galaxy is best described by the separate binary sub-types (e.g.: WD-WD, WD-NS, WD-BH, etc.) which make up the total population of binaries. Since it is quite likely that different sub-types will also have different population density descriptions, the most extensible and efficient way to describe the bulk properties of the Galaxy will be a list of parameters for each sub-type. The parameters will include the sub-type, a flag for the density distribution used, and the values of N (or ρ_0), c, R_0, and z_0.

The model of the Galaxy that is used for the MLDC is corresponds to Eq. (3.11) with $c = 1$, $R_0 = 2.5$ kpc, and $z_0 = 200$ pc. We have restricted our population to detached double white dwarf binaries. The total number of binaries in the population is $N = 3 \times 10^7$.

3.3.2 Stellar-mass binaries

Challenge-1 data sets include only signals from circular-orbit (i.e., eccentricity = 0) binaries, so we will concentrate on those. For circular-orbit SM binaries, the source parameters are

Galactic binaries source descriptors	
$\mathcal{A}$	overall signal amplitude
(θ, phi)	source position on the sky
(i, ψ)	inclination angle and polarization
f_0	GW frequency at time $t = 0$
φ_0	waveform phase $t = 0$ (beginning of data set)
$\dot{f}_0$	time-derivative of GW frequency at $t = 0$
$\ddot{f}_0$	second time-derivative of GW frequency at $t = 0$

Here $t = 0$ refers to time as measured at the SSB. The frequency derivative $\dot{f}$ can be due to either mass transfer or gravitational wave emission. For many sources $\dot{f}_0$ and/or $\ddot{f}_0$ are too small to be measured, and so can be left out of the source description. **In Challenge-1, for simplicity, we restrict to stellar-mass binaries that have circular orbits and constant frequency; i.e.,** $\dot{f}_0 = \ddot{f}_0 = 0$.

The amplitude $\mathcal{A}$ is given, in terms of the underlying physical parameters by

$$\mathcal{A} = \frac{2G^{5/3}}{c^4 D}(2\pi f)^{2/3}\mathcal{M}^{5/3}\,, \tag{3.12}$$

where D is the distance to the source, while the gravitational-wave frequency f is just

$$f = \frac{2}{P_{\rm orb}} \tag{3.13}$$

where $P_{\rm orb}$ is the orbital period. For binaries with constant frequency, $\mathcal{A}$ is clearly also a constant.

Within the quadrupole approximation, the amplitudes of the two polarizations are given by

$$A_+ = \mathcal{A}\left(1 + \cos^2\iota\right) \tag{3.14}$$

$$A_\times = -2\mathcal{A}\cos\iota. \tag{3.15}$$

$$\tag{3.16}$$

Full details concerning the stellar-binary waveforms can be found in Rubbo, Cornish & Poujade (2004) or Pierro *et al.* (2001).

3.4 Massive black hole binaries

A MBH binary is described by the two masses, sky location and distance, and (at any instant) the orbital angular frequency ω, the orbital phase Φ, the direction $\hat{\mathbf{L}}_\mathbf{N} \sim \mathbf{r} \times \mathbf{v}$ of the orbital angular momentum, as well as the two spins $\mathbf{S_1} \equiv \chi_1 m_1^2 \hat{\mathbf{S}}_\mathbf{1}$ and $\mathbf{S_2} \equiv \chi_2 m_2^2 \hat{\mathbf{S}}_\mathbf{2}$. Here $\hat{\mathbf{S}}_\mathbf{1}, \hat{\mathbf{S}}_\mathbf{2}$ are unit vectors and $0 \leq \chi_{1,2} < 1$. Bold fonts denote 3-d vectors and hats denote the unit vectors.

In subsection 3.4.1 we give high-order post-Newtonian orbital evolution equations for circular-orbit binaries, including spin effects, and in 3.4.2 we give equations for the corresponding waveform in the source and in radiation frames, as defined in (21; 20; 22). For Challenge-1 we specialize to the case of nonspinning MBHs, and include PN corrections only up through P^2N order.

3.4.1 Orbital evolution

For the orbital angular frequency evolution we use the PN approximation, formulated as a simple Taylor expansion, with spin-orbital (1.5PN) and spin-spin (formal 2PN) terms as given in (23) eqns.(1-7) with $\hat{\theta} = \theta - 3/7\lambda = \frac{1039}{4620}$. More explicitly:

$$\frac{d\omega}{dt} = \frac{96}{5}\frac{\eta}{M^2}(M\omega)^{11/3}\{1 + 1PN + 1.5PN + SO + 2PN + SS + 2.5PN + 3PN + 3.5PN\} \quad (3.17)$$

where

$$\begin{aligned}
1PN &= -\frac{743+924\eta}{336}(M\omega)^{2/3} && (3.18)\\
1.5PN &= 4\pi(M\omega) && (3.19)\\
SO &= -\frac{1}{12}\sum_{i=1,2}\left[\chi_i\left(\hat{\mathbf{L}}_\mathbf{N}.\hat{\mathbf{S}}_\mathbf{i}\right)\left(113\frac{m_i^2}{M^2}+75\eta\right)\right](M\omega) && (3.20)\\
2PN &= \left(\frac{34103}{18144}+\frac{13661}{2016}\eta+\frac{59}{16}\eta^2\right)(M\omega)^{4/3} && (3.21)\\
SS &= -\frac{1}{48}\eta\chi_1\chi_2\left[247(\hat{\mathbf{S}}_\mathbf{1}.\hat{\mathbf{S}}_\mathbf{2}) - 721(\hat{\mathbf{L}}_\mathbf{N}.\hat{\mathbf{S}}_\mathbf{1})(\hat{\mathbf{L}}_\mathbf{N}.\hat{\mathbf{S}}_\mathbf{2})\right](M\omega)^{4/3} && (3.22)\\
2.5PN &= -\frac{1}{672}(4159+14532\eta)\pi(M\omega)^{5/3} && (3.23)\\
3PN &= \left[\left(\frac{16447322263}{139708800}-\frac{1712}{105}\gamma_E+\frac{16}{3}\pi^2\right)+\left(-\frac{273811877}{1088640}+\frac{451}{48}\pi^2-\frac{88}{3}\hat{\theta}\right)\eta\right.\\
&\quad \left.+\frac{541}{896}\eta^2-\frac{5605}{2592}\eta^3-\frac{856}{105}\ln\left(16(M\omega)^{2/3}\right)\right](M\omega)^2 && (3.24)\\
3.5PN &= \left(-\frac{4415}{4032}+\frac{661775}{12096}\eta+\frac{149789}{3024}\eta^2\right)\pi(M\omega)^{7/3}, && (3.25)
\end{aligned}$$

where γ_E is Euler's constant. Note that transformation $t \to t/M, \omega \to \omega M$, leads to eliminating the total mass. This implies that the waveform for different total masses can be obtained by a simple re-scaling. (Of course this is just a reflection of the fact that there is no fundamental length scale in general relativity.)

For precession of the spins and the Newtonian angular momentum $\mathbf{L}_N = \eta M^2 (M\omega)^{-1/3}\hat{\mathbf{L}}_\mathbf{N}$ we can use eqns.(4.17a-c) in (20) or equivalently eqns.(8-10) in (23):

$$\begin{aligned}
\frac{d\hat{\mathbf{S}}_\mathbf{1}}{dt} &= \frac{(M\omega)^2}{2M}\left\{\eta(M\omega)^{-1/3}\left(4+3\frac{m_2}{m_1}\right)\hat{\mathbf{L}}_\mathbf{N} + \chi_2\frac{m_2^2}{M^2}\left[\hat{\mathbf{S}}_\mathbf{2} - 3(\hat{\mathbf{S}}_\mathbf{2}\cdot\hat{\mathbf{L}}_\mathbf{N})\hat{\mathbf{L}}_\mathbf{N}\right]\right\}\times\hat{\mathbf{S}}_\mathbf{1} && (3.26)\\
\frac{d\hat{\mathbf{S}}_\mathbf{2}}{dt} &= \frac{(M\omega)^2}{2M}\left\{\eta(M\omega)^{-1/3}\left(4+3\frac{m_1}{m_2}\right)\hat{\mathbf{L}}_\mathbf{N} + \chi_1\frac{m_1^2}{M^2}\left[\hat{\mathbf{S}}_\mathbf{1} - 3(\hat{\mathbf{S}}_\mathbf{1}\cdot\hat{\mathbf{L}}_\mathbf{N})\hat{\mathbf{L}}_\mathbf{N}\right]\right\}\times\hat{\mathbf{S}}_\mathbf{2} && (3.27)\\
\frac{d\hat{\mathbf{L}}_\mathbf{N}}{dt} &= \mathbf{V}_{\mathbf{L_N}}\times\hat{\mathbf{L}}_\mathbf{N} \equiv \frac{(M\omega)^2}{2M}\left\{\left(4+3\frac{m_2}{m_1}\right)\chi_1\frac{m_1^2}{M^2}\hat{\mathbf{S}}_\mathbf{1} + \left(4+3\frac{m_1}{m_2}\right)\chi_2\frac{m_2^2}{M^2}\hat{\mathbf{S}}_\mathbf{2} -\right.\\
&\quad \left. 3(M\omega)^{1/3}\eta\chi_1\chi_2\left[(\hat{\mathbf{S}}_\mathbf{2}\cdot\hat{\mathbf{L}}_\mathbf{N})\hat{\mathbf{S}}_\mathbf{1} + (\hat{\mathbf{S}}_\mathbf{1}\cdot\hat{\mathbf{L}}_\mathbf{N})\hat{\mathbf{S}}_\mathbf{2}\right]\right\}\times\hat{\mathbf{L}}_\mathbf{N}. && (3.28)
\end{aligned}$$

Alternatively, instead of the last equation we can use

$$\hat{\mathbf{L}}_\mathbf{N} = \{\sin(i)\cos(\alpha), \sin(i)\sin(\alpha), \cos(i)\} \quad (3.29)$$

along with the following evolution equations for α, i:

$$\frac{di}{dt} = V^y_{L_N}\cos(\alpha) - V^x_{L_N}\sin(\alpha) \tag{3.30}$$

$$\frac{d\alpha}{dt} = V^z_{L_N} - \frac{\cos(i)}{\sin(i)}\left[V^x_{L_N}\cos(\alpha) + V^y_{L_N}\sin(\alpha)\right] \tag{3.31}$$

where $\mathbf{V_{L_N}} = \{V^x_{L_N}, V^y_{L_N}, V^z_{L_N}\}$. In the implementation of this scheme (described below), we integrate equations for (i, α) and eqn.(3.28) and check consistency (error of integration).

Finally, for the phase evolution we have:

$$\frac{d\Phi}{dt} = \omega - \frac{d\alpha}{dt}\cos(i). \tag{3.32}$$

Besides masses M_1, M_2 and dimensionless spin magnitudes χ_1, χ_2, one has to specify the initial conditions. Initial data for the differential equations are specified at $t = t_0$. These are the initial directions of spins, $\hat{\mathbf{S}}_\mathbf{1}(t_0), \hat{\mathbf{S}}_\mathbf{2}(t_0)$, the initial direction of the orbital angular momentum $i(t_0), \alpha(t_0)$, the initial frequency $\omega_0 = \omega(t_0)$ and initial phase $\Phi_0 = \Phi(t_0)$. These values are defined in the source coordinate frame (see the figure in (22)).

Following (22), the evolution is terminated at

$$min_\omega \left\{\dot{\omega} = 0; \frac{dE_{3PN}}{d\omega} = 0\right\}, \tag{3.33}$$

which point is referred to as the MECO. One can write $\frac{dE_{3PN}}{d\omega} = 0$ explicitly using the above equations and

$$\frac{d}{d\omega} = \frac{1}{\dot{\omega}}\frac{d}{dt} \tag{3.34}$$

when we need to differentiate vectors. In implementing this prescription, we use $\dot{\omega}$ in the above equation only up to 1PN order, neglecting higher-order terms. Following this prescription and neglecting terms of cubic order or higher in the spins, one arrives at the following:

$$\begin{aligned}\frac{1}{M}\frac{dE_{3PN}}{d\omega} = -\frac{\mu}{3}(M\omega)^{-1/3}\Bigg\{1 - \frac{9+\eta}{6}(M\omega)^{2/3} + \frac{20}{3M^2}(\hat{\mathbf{L}}_\mathbf{N}.\mathbf{S_{eff}})(M\omega) + \\ \frac{5}{64}\left(1 - \frac{3}{8}\eta\right)(M\omega)^{1/3}\frac{\delta m}{M}\chi_1\chi_2\left[1 + \frac{743+924\eta}{336}(M\omega)^{2/3}\right](\hat{\mathbf{S}}_\mathbf{1}\times\hat{\mathbf{S}}_\mathbf{2})\hat{\mathbf{L}}_\mathbf{N} + \\ \frac{1}{8}(-81 + 57\eta - \eta^2)(M\omega)^{4/3} + \frac{3\eta}{4}\chi_1\chi_2\left[(\hat{\mathbf{S}}_\mathbf{1}.\hat{\mathbf{S}}_\mathbf{2}) - 3(\hat{\mathbf{L}}_\mathbf{N}.\hat{\mathbf{S}}_\mathbf{1})(\hat{\mathbf{L}}_\mathbf{N}.\hat{\mathbf{S}}_\mathbf{2})\right](M\omega)^{4/3} + \\ 4\left[-\frac{675}{64} + \left(\frac{34445}{576} - \frac{205}{96}\pi^2\right)\eta - \frac{155}{96}\eta^2 - \frac{35}{5184}\eta^3\right](M\omega)^2\Bigg\},\end{aligned} \tag{3.35}$$

where

$$\mathbf{S_{eff}} = \left(1 + \frac{3}{4}\frac{m_2}{m_1}\right)\mathbf{S_1} + \left(1 + \frac{3}{4}\frac{m_1}{m_2}\right)\mathbf{S_2} \tag{3.36}$$

For non-spinning binaries $\omega_{MECO} = 0.129M^{-1}$. Note this is almost twice as high as the frequency of the last stable orbit (LSO) in the test mass case ($\eta \to 0$): $\omega_{LSO} = 0.068M^{-1}$.

3.4.2 SMBH inspiral waveform

The inspiral waveform is defined in (20), eqn.(4.8) up to 1.5PN. and it is given explicitly in the Appendix B of (20) up to 1PN. In the radiation coordinate frame (see (21) eqns.(3.1-3.3), (20) eqns.(4.22a-c), (22) eqns.(21-23)) the waveform is given by

$$h^{ij}_{TT} = h_+ T^{ij}_+ + h_\times T^{ij}_\times \tag{3.37}$$

$$\mathbf{T}_+ = \mathbf{e}^R_x \otimes \mathbf{e}^R_x - \mathbf{e}^R_y \otimes \mathbf{e}^R_y \tag{3.38}$$

$$\mathbf{T}_\times = \mathbf{e}^R_x \otimes \mathbf{e}^R_y + \mathbf{e}^R_y \otimes \mathbf{e}^R_x \tag{3.39}$$

$$h_+ = \frac{1}{2} h^{ij} (T_+)_{ij}, \;\; h_\times = \frac{1}{2} h^{ij} (T_\times)_{ij} \tag{3.40}$$

The radiative system introduces one more angle, θ, which is defined as the inclination of the direction to the observer $\hat{\mathbf{N}}$ to the total angular momentum at t_0[1]. (We note that in this section $\hat{N}$ points from the source to observer, while in other sections we have used $\hat{n}$ as the unit vector pointing from observer to source, so $\hat{N} = -\hat{n}$.)

Finally the waveform (plus and cross polarizations in the radiation coordinate frame) can be written as

$$h_{+,\times} = \frac{2\mu}{D}\frac{M}{r}\left[Q_{+,\times} + \left(\frac{M}{r}\right)^{1/2} Q^1_{+,\times} + \left(\frac{M}{r}\right) Q^2_{+,\times} + \mathcal{O}\left(\frac{M}{r}\right)^{3/2} \right] \tag{3.41}$$

where expressions for $Q^1_{+,\times} = P^{0.5}_{+,\times}$, $Q^2_{+,\times} = P Q_{+,\times}$ are given in AppendixB of (20). In order to show the conventions used here we will give the formulae explicitly:

$$Q_{+,\times} = \frac{1}{2} Q^{ij} (T_{+,\times})_{ij}; \quad Q^1_{+,\times} = \frac{1}{2} (Q^1)^{ij} (T_{+,\times})_{ij}; \quad Q^2_{+,\times} = \frac{1}{2} (Q^2)^{ij} (T_{+,\times})_{ij}. \tag{3.42}$$

where

$$Q^{ij} = 2(\hat{v}^i \hat{v}^j - \hat{r}^i \hat{r}^j), \tag{3.43}$$

$$(Q^1)^{ij} = \frac{\delta m}{m}\left\{ 6(\hat{\mathbf{N}}\cdot\hat{\mathbf{r}})\hat{r}^{(i}\hat{v}^{j)} + (\hat{\mathbf{N}}\cdot\hat{\mathbf{v}})(\hat{r}^i\hat{r}^j - 2\hat{v}^i\hat{v}^j) \right\} \tag{3.44}$$

$$\begin{aligned}(Q^2)^{ij} = {} & \frac{2}{3}(1-3\eta)\left\{ (\hat{\mathbf{N}}\cdot\hat{\mathbf{r}})^2(5\hat{r}^i\hat{r}^j - 7\hat{v}^i\hat{v}^j) - 16(\hat{\mathbf{N}}\cdot\hat{\mathbf{r}})(\hat{\mathbf{N}}\cdot\hat{\mathbf{v}})\hat{r}^{(i}\hat{v}^{j)} + (\hat{\mathbf{N}}\cdot\hat{\mathbf{v}})^2(3\hat{v}^i\hat{v}^j - \hat{r}^i\hat{r}^j) \right\} \\ & \frac{1}{3}(19-3\eta)(\hat{r}^i\hat{r}^j - \hat{v}^i\hat{v}^j) + \frac{2}{M^2} n^{(i}(\boldsymbol{\Delta}\times\hat{\mathbf{N}})^{j)},\end{aligned} \tag{3.45}$$

where $\hat{\mathbf{v}}$ is the unit vector along the relative velocity $\mathbf{v}$, $\hat{\mathbf{r}}$ is unit vector along the separation vector of the binary $\mathbf{r}$, $\hat{\mathbf{N}}$ is the unit vector that points from source to observer, and

$$\frac{\boldsymbol{\Delta}}{M^2} \equiv \chi_2 \frac{m_2}{M}\hat{\mathbf{S}}_\mathbf{2} - \chi_1 \frac{m_1}{M}\hat{\mathbf{S}}_\mathbf{1}. \tag{3.46}$$

In the source coordinate frame:

$$\begin{aligned}\hat{\mathbf{r}} &= \{-\sin(\alpha)\cos(\Phi) - \cos(\alpha)\cos(i)\sin(\Phi), \cos(\alpha)\cos(\Phi) - \sin(\alpha)\cos(i)\sin(\Phi), \sin(i)\sin(\Phi)\} \\ \hat{\mathbf{v}} &= \{\sin(\alpha)\sin(\Phi) - \cos(\alpha)\cos(i)\cos(\Phi), -\cos(\alpha)\sin(\Phi) - \sin(\alpha)\cos(i)\cos(\Phi), \sin(i)\cos(\Phi)\}, \\ \hat{\mathbf{N}} &= \{\sin(\theta), 0, \cos(\theta)\}.\end{aligned} \tag{3.47}$$

[1] Another angle, ϕ, defining the direction to the observer, can always be put to zero by rotating the coordinate frame around the z-axis

The radiation frame is related to the source frame in the following way:

$$\hat{\mathbf{e}}^{\mathbf{R}}_{\mathbf{x}} = \hat{\mathbf{e}}^{\mathbf{S}}_{\mathbf{x}}\cos(\theta) - \hat{\mathbf{e}}^{\mathbf{S}}_{\mathbf{z}}\sin(\theta) \tag{3.48}$$

$$\hat{\mathbf{e}}^{\mathbf{R}}_{\mathbf{y}} = \hat{\mathbf{e}}^{\mathbf{S}}_{\mathbf{y}} \tag{3.49}$$

$$\hat{\mathbf{e}}^{\mathbf{R}}_{\mathbf{z}} = \hat{\mathbf{e}}^{\mathbf{S}}_{\mathbf{x}}\sin(\theta) + \hat{\mathbf{e}}^{\mathbf{S}}_{\mathbf{z}}\cos(\theta) = \hat{\mathbf{N}}. \tag{3.50}$$

So that

$$\mathbf{T}_{+} = \hat{\mathbf{e}}^{\mathbf{S}}_{\mathbf{x}} \otimes \hat{\mathbf{e}}^{\mathbf{S}}_{\mathbf{x}}\cos^2(\theta) - \hat{\mathbf{e}}^{\mathbf{S}}_{\mathbf{x}} \otimes \hat{\mathbf{e}}^{\mathbf{S}}_{\mathbf{z}}\sin(\theta)\cos(\theta) - \hat{\mathbf{e}}^{\mathbf{S}}_{\mathbf{z}} \otimes \hat{\mathbf{e}}^{\mathbf{S}}_{\mathbf{x}}\sin(\theta)\cos(\theta) - \hat{\mathbf{e}}^{\mathbf{S}}_{\mathbf{y}} \otimes \hat{\mathbf{e}}^{\mathbf{S}}_{\mathbf{y}} + \hat{\mathbf{e}}^{\mathbf{S}}_{\mathbf{z}} \otimes \hat{\mathbf{e}}^{\mathbf{S}}_{\mathbf{z}}\sin^2(\theta) \tag{3.51}$$

$$\mathbf{T}_{\mathbf{x}} = (\hat{\mathbf{e}}^{\mathbf{S}}_{\mathbf{x}} \otimes \hat{\mathbf{e}}^{\mathbf{S}}_{\mathbf{y}} + \hat{\mathbf{e}}^{\mathbf{S}}_{\mathbf{y}} \otimes \hat{\mathbf{e}}^{\mathbf{S}}_{\mathbf{x}})\cos(\theta) - (\hat{\mathbf{e}}^{\mathbf{S}}_{\mathbf{y}} \otimes \hat{\mathbf{e}}^{\mathbf{S}}_{\mathbf{z}} + \hat{\mathbf{e}}^{\mathbf{S}}_{\mathbf{z}} \otimes \hat{\mathbf{e}}^{\mathbf{S}}_{\mathbf{y}})\sin(\theta) \tag{3.52}$$

Using these relationships we have for the Qs (we give the expressions for "+" polarization only, "x" polarization can be obtained by replacing "+" with "x"):

$$Q_{+} = -2(C_{+}\cos(2\Phi) + S_{+}\sin(2\Phi)) \tag{3.53}$$

$$\begin{aligned} Q^1_{+} = {} & \frac{1}{4}\frac{\delta m}{M}\{9(aS_{+} + bC_{+})\cos(3\Phi) + 9(bS_{+} - aC_{+})\sin(3\Phi) + \\ & (3aS_{+} - 3bC_{+} - 2bK+)\cos(\Phi) - (3bS_{+} + 3aC_{+} - 2aK_{+})\sin(\Phi)\} \end{aligned} \tag{3.54}$$

$$\begin{aligned} Q^2_{+} = {} & \frac{8}{3}(1-3\eta)\left\{[(a^2-b^2)C_{+} - 2abS_{+}]\cos(4\Phi) + [(a^2-b^2)S_{+} + 2abC_{+}]\sin(4\Phi)\right\} \\ & + \frac{1}{3}\left\{2(1-3\eta)[(b^2-a^2)K_{+} + 2(a^2+b^2)C_{+}] + (19-3\eta)C_{+}\right\}\cos(2\Phi) \\ & + \frac{1}{3}\left\{4(1-3\eta)[(a^2+b^2)S_{+} - abK_{+}] + (19-3\eta)S_{+}\right\}\sin(2\Phi) \\ & + DC_{+}\cos(\Phi) + DS_{+}\sin(\Phi). \end{aligned} \tag{3.55}$$

where

$$
\begin{aligned}
C_+ &= \frac{1}{2}\cos^2(\theta)\left(\sin^2(\alpha)-\cos^2(i)\cos^2(\alpha)\right)+\frac{1}{2}\left(\cos^2(i)\sin^2(\alpha)-\cos^2(\alpha)\right)- \\
&\quad \frac{1}{2}\sin^2(\theta)\sin^2(i)-\frac{1}{4}\sin(2\theta)\sin(2i)\cos(\alpha), && (3.56)\\
C_\times &= -\frac{1}{2}\cos(\theta)\sin(2\alpha)\left(1+\cos^2(i)\right)-\frac{1}{2}\sin(\theta)\sin(2i)\sin(\alpha), && (3.57)\\
S_+ &= \frac{1}{2}\left(1+\cos^2(\theta)\right)\cos(i)\sin(2\alpha)+\frac{1}{2}\sin(2\theta)\sin(i)\sin(\alpha), && (3.58)\\
S_\times &= -\cos(\theta)\cos(i)\cos(2\alpha)-\sin(\theta)\sin(i)\cos(\alpha) && (3.59)\\
K_+ &= \frac{1}{2}\cos^2(\theta)\left(\sin^2(\alpha)+\cos^2(i)\cos^2(\alpha)\right)-\frac{1}{2}\left(\cos^2(i)\sin^2(\alpha)+\cos^2(\alpha)\right)+ \\
&\quad \frac{1}{2}\sin^2(\theta)\sin^2(i)+\frac{1}{4}\sin(2\theta)\sin(2i)\cos(\alpha), && (3.60)\\
K_\times &= -\frac{1}{2}\cos(\theta)\sin(2\alpha)\sin^2(i)+\frac{1}{2}\sin(\theta)\sin(2i)\sin(\alpha), && (3.61)\\
DC_+ &= -\frac{1}{M^2}\left[\Delta^y\sin(\alpha)\cos(\theta)+d\cos(\alpha)\right], && (3.62)\\
DS_+ &= -\frac{1}{M^2}\left[c\Delta^y-d\cos(i)\sin(\alpha)\right], && (3.63)\\
DC_\times &= \frac{1}{M^2}\left[\Delta^y\cos(\alpha)-d\cos(\theta)\sin(\alpha)\right], && (3.64)\\
DS_\times &= \frac{1}{M^2}\left[-\Delta^y\cos(i)\sin(\alpha)+cd\right], && (3.65)\\
a &= -\sin(\theta)\sin(\alpha), && (3.66)\\
b &= \cos(\theta)\sin(i)-\sin(\theta)\cos(i)\cos(\alpha), && (3.67)\\
c &= \cos(\theta)\cos(i)\cos(\alpha)+\sin(i)\sin(\theta), && (3.68)\\
d &= \Delta^z\sin(\theta)-\Delta^x\cos(\theta). && (3.69)
\end{aligned}
$$

Since we integrate with respect to the orbital angular frequency, we need to relate M/r to $(M\omega)$. With sufficient accuracy for our purpose, this is given by:

$$\frac{M}{r}\approx(M\omega)^{2/3}\left[1+\frac{1}{3}(3-\eta)(M\omega)^{2/3}\right]$$

Taking the above into account we can rewrite $h_{+,\times}$ as follows:

$$h_{+,\times}=\frac{2\mu}{D}(M\omega)^{2/3}\left[Q_{+,\times}+(M\omega)^{1/3}Q^1_{+,\times}+(M\omega)^{2/3}\left(Q^2_{+,\times}+\frac{1}{3}(3-\eta)Q_{+,\times}\right)+\mathcal{O}(M\omega)\right] \quad (3.70)$$

The last term in (3.70) yields the following change in (3.55): instead of $(19-3\eta)C_+$, $(19-3\eta)S_+$ we have $(13-\eta)C_+$, $(13-\eta)S_+$ correspondingly.

In order to generate waveform using the inspiralling trajectory described in the previous subsection one need to specify distance between observer and the source, and the direction θ to the observer in the source frame. (Again, we can always choose $\phi=0$.)

3.4.3 Reduction to the non-spinning case

Reduction from the spinning to the non-spinning case is quite simple. We start with the orbital evolution.

3.4.4 Orbital evolution: non-spinning case

The orbital evolution is described by only one differential equation, which can be integrated analytically. This is the equation for the orbital frequency (3.17) with $SO = SS = 0$. The analytic expressions for frequency and phase are given in (24); we also give them here. First introduce

$$\tau = \frac{\eta}{5M}(t_c - t) \tag{3.71}$$

$$x = v^2 = (M\omega)^{2/3} \tag{3.72}$$

Then

$$\begin{aligned}(M\omega)^{2/3} &= \frac{1}{4}\tau^{-1/4}\left\{1 + \left(\frac{743}{4032} + \frac{11}{48}\eta\right)\tau^{-1/4} - \frac{1}{5}\pi\tau^{-3/8} + \right. \\ &\left(\frac{19583}{254016} + \frac{24401}{193536}\eta + \frac{31}{288}\eta^2\right)\tau^{-1/2} + \left(-\frac{11891}{53760} + \frac{109}{1920}\eta\right)\pi\tau^{-5/8} - \\ &\left[-\frac{10052469856691}{6008596070400} + \frac{1}{6}\pi^2 + \frac{107}{420}\gamma_E - \frac{107}{3360}\ln\left(\frac{\tau}{256}\right) + \left(\frac{15335597827}{3901685760} - \frac{451}{3072}\pi^2 - \right.\right. \\ &\left.\left.\frac{77}{72}\lambda + \frac{11}{24}\theta\right)\eta - \frac{15211}{442368}\eta^2 + \frac{25565}{331776}\eta^3\right]\tau^{-3/4} + \left(-\frac{113868647}{433520640} - \frac{31821}{143360}\eta + \right. \\ &\left.\left.\frac{294941}{3870720}\eta^2\right)\pi\tau^{-7/8}\right\}\end{aligned} \tag{3.73}$$

where $\lambda = -\frac{1987}{3080}, \theta = -\frac{11831}{9240}$. Or taking the power 3/2:

$$\begin{aligned}M\omega &= \frac{1}{8}\tau^{-3/8}\left\{1 + \left(\frac{11}{32}\eta + \frac{743}{2688}\right)\tau^{-1/4} - \frac{3}{10}\pi\tau^{-3/8} + \left(\frac{1855099}{14450688} + \frac{371}{2048}\eta^2 + \frac{56975}{258048}\eta\right)\tau^{-1/2} + \right. \\ &\left(\frac{13}{256}\eta - \frac{7729}{21504}\right)\pi\tau^{-5/8} + \left[\frac{235925}{176472}\eta^3 - \frac{30913}{1835008}\eta^2 + \left(-\frac{451}{2048}\pi^2 + \frac{25302017977}{4161798144}\right)\eta - \right. \\ &\left.\frac{720817631400877}{288412611379200} + \frac{53}{200}\pi^2 + \frac{107}{280}\gamma_E - \frac{107}{2240}\ln\left(\frac{\tau}{256}\right)\right]\tau^{-3/4} + \\ &\left.\left(\frac{141769}{1290240}\eta^2 - \frac{188516689}{433520640} - \frac{97765}{258048}\eta\right)\pi\tau^{-7/8}\right\}\end{aligned} \tag{3.74}$$

The orbital phase can be expressed in terms of orbital frequency as follows:

$$\begin{aligned}\Phi(\omega) &= -\frac{1}{32\eta}(M\omega)^{-5/3}\left\{1 + \left(\frac{3715}{1008} + \frac{55}{12}\eta\right)(M\omega)^{2/3} - 10\pi(M\omega) + \right. \\ &\left(\frac{15293365}{1016064} + \frac{27145}{1008}\eta + \frac{3085}{144}\eta^2\right)(M\omega)^{4/3} + \left(\frac{38645}{1344} - \frac{65}{16}\eta\right)\pi\ln\left(\frac{\omega}{\omega_0}\right)(M\omega)^{5/3} + \\ &\left[\frac{12348611926451}{18776862720} - \frac{160}{3}\pi^2 - \frac{1712}{21}\gamma_E - \frac{856}{21}\ln\left[16(m\omega)^{2/3}\right] + \left(-\frac{15335597827}{12192768} + \frac{2255}{48}\pi^2 + \frac{3080}{9}\lambda - \right.\right. \\ &\left.\left.\frac{440}{3}\theta\right)\eta - \frac{76055}{6912}\eta^2 - \frac{127825}{5184}\eta^3\right](M\omega)^2 + \left(\frac{77096675}{2032128} + \frac{378515}{12096}\eta - \frac{74045}{6048}\eta^2\right)\pi(M\omega)^{7/3}\right\}\end{aligned} \tag{3.75}$$

This is a "TaylorT3" waveform in the classification scheme of (25), if one substitutes orbital angular frequency (3.74) in the expression for Φ above.

3.4.5 Gravitational waveform for the non-spinning case

The reduction of the waveform to the non-spinning is easily accomplished by using (3.70) and setting

$$\alpha = 0, \quad i = 0. \tag{3.76}$$

together with $\chi_1 = \chi_2 = 0$. This reduces the long expression to a relatively short one [2]:

$$C_+ = -\frac{1}{2}(1 + \cos^2(\theta)), \quad C_\times = 0 \tag{3.77}$$

$$S_+ = 0, \quad S_\times = -\cos(\theta) \tag{3.78}$$

$$K_+ = -\frac{1}{2}\sin^2(\theta), \quad K_\times = 0 \tag{3.79}$$

$$DC_+ = DC_\times = DS_+ = DS_\times = 0 \tag{3.80}$$

$$a = 0, \quad d = 0, \quad b = -\sin(\theta), \quad c = \cos(\theta). \tag{3.81}$$

Substituting these expressions into (3.55) we get:

$$Q_+ = (1 + \cos^2(\theta))\cos(2\Phi) \tag{3.82}$$

$$Q_\times = 2\cos(\theta)\sin(2\Phi) \tag{3.83}$$

$$Q_+^{(1)} = \frac{1}{8}\frac{\delta m}{M}\sin(\theta)\left[9(1 + \cos^2(\theta))\cos(3\Phi) - (5 + \cos^2(\theta))\cos(\Phi)\right] \tag{3.84}$$

$$Q_\times^{(1)} = \frac{3}{4}\frac{\delta m}{M}\sin(\theta)\cos(\theta)\left[3\sin(3\Phi) - \sin(\Phi)\right] \tag{3.85}$$

$$\begin{aligned}\frac{1}{3}(3-\eta)Q_+ + Q_+^{(2)} = {} & \frac{4}{3}(1-3\eta)(1-\cos^4(\theta))\cos(4\Phi) + \\ & \frac{1}{6}\left[-19 + 19\eta - (9 + 11\eta)\cos^2(\theta) + (1-3\eta)\cos^4(\theta)\right]\cos(2\Phi)\end{aligned} \tag{3.86}$$

$$\begin{aligned}\frac{1}{3}(3-\eta)Q_\times + Q_\times^{(2)} = {} & \frac{8}{3}(1-3\eta)(1-\cos^2(\theta))\cos(\theta)\sin(4\Phi) - \\ & \frac{1}{3}\left[17 - 13\eta - 4(1-3\eta)\cos^2(\theta)\right]\cos(\theta)\sin(2\Phi)\end{aligned} \tag{3.87}$$

which is in agreement with the result derived by a number of authors.

3.4.6 Further simplification of the SMBH signal model for Challenge-1 mock data

For Challenge-1 data sets, we further simplify the signal model. Not only are spins set to zero, but also in the waveform phase we neglect PN corrections above P^2N order. Also we neglect all harmonics except the dominant $m = 2$ (mass-quadrupole) piece, and we neglect PN corrections to the amplitude of that harmonic. (These last 2 simplications are sometimes referred to as the "restricted" PN approximation.) The Challenge-1 waveforms are therefore especially simple. The evolution of the orbital angular frequency and orbital phase up to 2PN order are given by:

$$\begin{aligned}M\omega = {} & \frac{1}{8}\tau^{-3/8}\left\{1 + \left(\frac{11}{32}\eta + \frac{743}{2688}\right)\tau^{-1/4} - \frac{3}{10}\pi\tau^{-3/8}\right. \\ & + \left.\left(\frac{1855099}{14450688} + \frac{371}{2048}\eta^2 + \frac{56975}{258048}\eta\right)\tau^{-1/2}\right\}\end{aligned} \tag{3.88}$$

[2] Note that here the angle θ corresponds to the angle call i in (21)

λ^0	M_1	first mass
λ^1	M_2	second mass
λ^2	ω_0	orbital angular frequency
λ^3	Φ_0	orbital initial phase
λ^4	θ	inclination of L (orbital angular momentum) to the direction to the observer
λ^5	D	luminocity distance
λ^6	ψ	polarization angle, required only to present the signal in SSB frame

Table 3.1: Summary of physical parameters for BBH and their meaning. The orbital angular frequency is given at the beginning of observation which is taken as zero-time.

And the phase is described by:

$$\begin{aligned}\Phi &= -\frac{1}{32\eta}(M\omega)^{-5/3}\left\{1+\left(\frac{3715}{1008}+\frac{55}{12}\eta\right)(M\omega)^{2/3}-10\pi(M\omega)\right.\\ &+ \left.\left(\frac{15293365}{1016064}+\frac{27145}{1008}\eta+\frac{3085}{144}\eta^2\right)(M\omega)^{4/3}\right\} \end{aligned} \tag{3.89}$$

To terminate the inspiral evolution, we could imagine using any of the three following criteria. The first is to terminate at $r = 6M$, which corresponds to

$$\omega_{LSO} = 0.0138(10^6 M_\odot/M)Hz \tag{3.90}$$

Alternatively one can use the "MECO" condition at 2PN:

$$\omega_{MECO} = \frac{1}{27}\left[-\frac{54+6\eta-6\sqrt{1539-1008\eta+19\eta^2}}{81-57\eta+\eta^2}\right]\frac{M_\odot}{M}\,Hz \tag{3.91}$$

and/or condition the condition.

$$\dot{\omega} > 0.$$

In practice we use whichever is smaller (i.e., we terminate the signal as soon as one of those conditions is met).

We also note that for some Challenge data sets we artificially terminate the signal a week or so before the coalescence time t_c. The goal there is to approximately reproduce a situation where it is still roughly one week before merger, but one wants to alert astronomers where in the sky to point their telescopes (and at what precise time) in order to perhaps catch an electromagnetic counterpart to the GW signal. The goal of that type of Challenge is to demonstrate how well that can be done, in practice.

The waveform is described by its two polarizations in the radiation frame:

$$h_+ = \frac{2\mu}{D}(M\omega)^{2/3}(1+\cos^2(\theta))\cos(2\Phi) \tag{3.92}$$

$$h_\times = \frac{4\mu}{D}(M\omega)^{2/3}\cos(\theta)\sin(2\Phi) \tag{3.93}$$

Since we terminate waveform suddenly in time domain and signal spectrum has a large dynamical range, we have spectral leakage in the frequency domain, which we deal with by applying

window (tapering) in the time domain. We have chosen the following taper, which causes exponential damping of the signal in time domain after time $t_R = t(R)$, time when radius of orbit is R:

$$w(t) = \frac{1}{2}\left(1 + \tanh\left[A(t_k - t)\right]\right) \tag{3.94}$$

where $t_k = (t_c + t_R)/2$ and coefficient $A = 0.1\left(\frac{7M}{R}\right)^{7.5}$ was found empirically to produce smooth damping for R in range $[7M, 9M]$.

The transformation to the SSB system is given in the subsection 3.2.3:

$$h_+^{SSB}(\xi) = \cos(2\psi)\, h_+(\xi) + \sin(2\psi)\, h_\times(\xi), \tag{3.95}$$

$$h_\times^{SSB}(\xi) = \cos(2\psi)\, h_\times(\xi) - \sin(2\psi)\, h_+(\xi). \tag{3.96}$$

This is finalizes the time domain description of the waveform, which is the input to *Synthetic LISA* and/or the *LISA Simulator*.

3.4.7 Implementation

The waveform generation is implemented as C++ class `BBHChallenge1`

- `SpinBBHWaveform(float mass1, float mass2)`, constructor, the input are two masses in units of solar mass.
- `SetInitialOrbit(float omega0, float phi0)`, defines initial orbit: `omega0` is initial angular orbital frequency and `phi0` is initial orbital phase
- `double EstimateTc(float omega0)`, this method is used to estimate coalescence time t_c for a given initial angular orbital frequency. It is done by numerical solution of eqn.(3.88).
- `double EstimateFreq0(float tc)`, this method computes initial angular orbital frequency for a given coalescence time.
- `ComputeInspiral(float timeStep, float maxDuration)` this method computes orbital evolution. Method `SetInitialOrbit` should be called beforehand.
- `SetObserver(float thetaD, double D)` this method initializes position of the observer with respect to the source frame: `thetaD` is θ (see, eqns.(3.92, 3.93)) and `D` is luminosity distance (in pc).
- `ComputeWaveform(float truncateTime, float taper, std::vector<double>& hPlus, std::vector<double>& hCross)`. This method computes waveform in the radiation frame, eqns.(3.92,3.93). The variable `truncateTime` was introduced to truncate the waveform before coalescence ("early warning") but it is not used in the current challenge (always set to 0), `taper` is used to terminate waveform smoothly (see explanations in the previous subsection), and it is hardcoded to $8M$ in the first round of challenges.
- `ComputeWaveformSSB(float truncateTime, float taper, float psi, std::vector<double>& hPlus, std::vector<double>& hCross)`. The same as above, but computes waveform in the SSB frame according to eqn.(3.95, 3.96).

3.4.8 Stationary phase approximation

Here we derive the Fourier transform of the time-domain signal described above using stationary phase approximation (SPA). The SPA was *not* used in creating the signals injected into Challenge-1 data sets, but it could be useful for search algorithms, so we give the main results here.

The signal in the radiation frame is given by eqns.(3.92, 3.93), this signal has to be passed through a particular transfer function to get a TDI observable, this can be presented in a rather symbolic form as follows

$$h^I = \sum_{j=1}^{3} \sum_k A^I_{k,j}(t) \left[F^+_j(t) h_+(t-t_k) + F^\times_j(t) h_\times(t-t_k) \right] \tag{3.97}$$

where index I corresponds to a different TDI combination α_i, $\bar{A}, \bar{E}, \bar{T}, ...$; t_k are delays and $A_{k,j}(t)$ scales an amplitude of each delayed waveform. The antenna pattern if given as

$$\begin{aligned} F^+_j &= u_j(t)\cos(2\psi) + v_j(t)\sin(2\psi), & (3.98)\\ F^\times_j &= v_j(t)\cos(2\psi) - u_j(t)\sin(2\psi) & (3.99) \end{aligned}$$

where ψ is polarization angle used in eqn.(3.95,3.96). The functions $u_j(t), v_j(t)$ depend on the rotation matrix O_1 and on LISA-to-SSB transformation, thus, they depend on time through motion of the guiding center and cartwheelng motion of the spacecrafts, and on the position of the source in the sky, given by the ecliptic coordinates β, λ. The explicit expressions for those functions are given in (28), eqns.(27-39). Following (26) introduce polarization phase $\Phi^j_p(t)$:

$$F^+_j(t) h_+(t-t_k) + F^\times_j(t) h_\times(t-t_k) = \mathcal{A}^j(t) (M\omega)^{2/3} \cos(2\Phi(t-t_k) - \Phi^j_p(t)) \tag{3.100}$$

where

$$\begin{aligned} \mathcal{A}^j(t) &= \frac{2\mu}{D}\sqrt{(F^+_j(t)h^+_0)^2 + (F^\times_j(t)h^\times_0)^2}, & (3.101)\\ \tan(\Phi^j_p) &= \frac{F^\times_j(t)h^\times_0}{F^+_j(t)h^+_0} & (3.102)\\ h^\times_0 &= 2\cos(\theta), \quad h^+_0 = \frac{1}{2}(1+\cos^2(\theta)) & (3.103) \end{aligned}$$

As a next step we expand the phase around t:

$$\Phi(t-t_k) \approx \Phi(t) - \omega t_k + ... \tag{3.104}$$

here we neglected second order Doppler corrections (see (26) for justification). It also convenient to represent cos in exponential form:

$$\cos(2\Phi(t-t_k) - \Phi^j_p(t)) = \frac{1}{2} e^{-i\left[2\Phi(t) - 2\omega t_k - \Phi^j_p(t)\right]} + c.c. \tag{3.105}$$

where *c.c.* stands for complex conjugate, which we will omit in further expressions (I assume that we mainly interested in positive frequencies). Here we can identify the Doppler phase $\Phi^k_D(t) = 2\omega t_k$. The delays can be further expanded according to:

$$t_j = R\cos(\beta)\cos(\Omega t + \eta_0 - \lambda) + \hat{\mathbf{k}}(O_2(t)\mathbf{p}^L_j) - m_j L \tag{3.106}$$

where $R = 1au$, η_0 is initial phase of the LISA's guiding center, O_2 is the LISA-to-SSB rotation matrix (see (28)), $\mathbf{p}_j^L$ is the vector connecting the guiding center and j-th spacecraft and m_j is integer coming from a particular TDI combination. Note that $\omega\hat{\mathbf{k}}(O_2(t)\mathbf{p}_j^L) \sim \omega L$.

Put all pieces together:

$$h^I = \Lambda^I(t)(M\omega)^{2/3}e^{-i(2\Phi(t))}, \tag{3.107}$$

$$\Lambda^I(t) = \sum_{j=1}^{3}\sum_{k}\frac{1}{2}A^I_{k,j}(t)\mathcal{A}^j(t)e^{i(\Phi^k_D(t)+\Phi^j_p(t))} \tag{3.108}$$

Since $\Lambda(t)$ varies on the timescale $1yr >> 2\pi/\omega$ we can apply SPA to evaluate the waveform in the frequency domain:

$$\tilde{h}^I_{spa} = \Lambda(t_f)(M\omega(t_f))^{2/3}\left[\frac{\pi}{\dot{\omega}(t_f)}\right]^{1/2}e^{i(2\pi f t_f - 2\Phi(t_f)-\pi/4)} \tag{3.109}$$

where t_f is defined by equation

$$\omega(t_f) = \pi f \tag{3.110}$$

For restricted waveform it is sufficient to use only leading order term in amplitude:

$$(\dot{\omega}(t_f))^{-1/2} = M\sqrt{\frac{5}{96\eta}}(\pi Mf)^{-11/6} \tag{3.111}$$

The phase (up to 2PN order) is given in (25):

$$\begin{aligned} 2\pi f t_f - 2\Phi(t_f) &= 2\pi f t_c - \phi_c + \frac{3}{128\eta}(\pi Mf)^{-5/3}\left\{1+\frac{5}{9}\left(\frac{743}{84}+11\eta\right)(\pi Mf)^{2/3} - 16\pi(\pi Mf) + \right. \\ &\left.\left(\frac{15293365}{1016064}+\frac{27145}{1008}\eta+\frac{3085}{144}\eta^2\right)(\pi Mf)^{4/3}\right\} \end{aligned} \tag{3.112}$$

and inverting (3.88) we obtain

$$\begin{aligned} t_f &= t_c - \frac{5M}{256\eta}(\pi Mf)^{-8/3}\left\{1+\left(\frac{743}{252}+\frac{11}{3}\eta\right)(\pi Mf)^{2/3} - \frac{32}{5}\pi(\pi Mf) + \right. \\ &\left.\left(\frac{8579163}{508032}+\frac{29555}{1008}\eta+\frac{203}{8}\eta^2\right)(\pi Mf)^{4/3}\right\} \end{aligned} \tag{3.113}$$

3.5 EMRIs: Analytic kludge waveforms for extreme-mass-ratio inspirals

3.5.1 Introduction

The EMRI waveforms injected into Challenge-1 data sets are the "analytic kludge" version of these waveforms, introduced by Barack & Cutler (29), hereafter referred to as BC. In the analytic kludge, the orbit is approximated as a Newtonian ellipse at any instant, but one whose perihelion direction, orbital plane, semi-major axis, and eccenticity evolve according to post-Newtonian evolution equations. At any instant, the emitted waveform is taken to be the Peters-Matthews waveform corresponding to the instantaneous Newtonian orbit. While these waveforms are not particularly

accurate in the highly relativistic regime of interest for EMRI searches, they do exhibit the main qualitative features of true waveforms, and are considerably simpler to generate. It is expected that any search strategy that works for "analytic kludge" waveforms could be converted over fairly easily to true, GR waveforms, once these become available. For more details we refer the reader to BC (from which pieces were chopped up and re-pasted to make this section).

Our index notation in this section is the following. Indices for vectors and tensors on parameter space are chosen from the beginning of the Latin alphabet $(a, b, c, \ldots)$. Vectors and tensors on three-dimensional space have indices chosen from the middle of the Latin alphabet $(i, j, k, \ldots)$, and run over $1, 2, 3$; their indices are raised and lowered with the flat 3-metric, η_{ij}. Actually, we adopt a mixed notation for spatial vectors, sometimes labelling them with spatial indices $(i, j, k, \ldots)$, but sometimes suppressing the indices and instead using the standard $3-d$ vector notation: an over-arrow (as in $\vec{A}$) to represent a vector, $\vec{A}\cdot\vec{B}$ to represent a scalar ("dot") product, and $\vec{A}\times\vec{B}$ to represent the vector ("cross") product. An over-hat (as in $\hat{n}$) will indicate that a vector is normalized, i.e., has unit length. We trust our meaning will always be clear, despite this mixed notation.

3.5.2 Parameter space

The two-body system is described by 17 parameters. The spin of the CO can be marginally relevant (see Appendix C of BC), but in this paper we shall ignore it, leaving us with 14 parameters. We shall denote a vector in the 14-d parameter space by λ^a $(a = 0, \ldots, 13)$. We choose our parameters as follows:

$$\begin{aligned}\lambda^a \;\equiv\; & (\lambda^0, \ldots, \lambda^{13}) = \\ & \left[t_0\ \ln\mu,\ \ln M,\ S/M^2,\ e_0,\ \tilde{\gamma}_0,\ \Phi_0,\ \mu_S \equiv \cos\theta_S,\ \phi_S,\ \cos\lambda,\ \alpha_0, \mu_K \equiv \cos\theta_K,\ \phi_K,\ \ln(\mu/M),\ D\right]\end{aligned} \tag{3.114}$$

Here, t_0 is a time parameter that allows us to specify "when" the inspiral occurs—we shall generally choose t_0 to be the instant of time when the (radial) orbital frequency sweeps through some fiducial value ν_0 (typically, we shall choose ν_0 of order 1 mHz), μ and M are the masses of the CO and MBH, respectively, and S is the magnitude of the MBH's spin angular momentum (so $0 \le S/M^2 \le 1$). The parameters e_0, $\tilde{\gamma}_0$, and Φ_0 describe, respectively, the eccentricity, the direction of the pericenter within the orbital plane, and the mean anomaly—all at time t_0. More specifically, we take $\tilde{\gamma}_0$ to be the angle (in the plane of the orbit) from $\hat{L}\times\hat{S}$ to pericenter, and, as usual, Φ_0 to be the mean anomaly with respect to pericenter passage. The parameter $\alpha_0 \equiv \alpha(t = t_0)$ [where $\alpha(t)$ is defined in Eq. (3.127)] describes the direction of $\hat{L}$ around $\hat{S}$ at t_0. The angles (θ_S, ϕ_S) are the direction to the source, in ecliptic-based coordinates; $(\theta_K,\ \phi_K)$ represent the direction $\hat{S}$ of the MBH's spin (approximated as constant) in ecliptic-based coordinates; and λ is the angle between $\hat{L}$ and $\hat{S}$ (also approximated as constant[3]). Finally, D is the distance to the source.

The various parameters and their meaning are summarized in Table 3.2. Fig. 3.1 illustrates the various angles involved in our parameterization.

The parameters can be divided into "intrinsic" and "extrinsic" parameters, following Buonanno, Chen, and Vallisneri (34) (hereafter, BCV). Extrinsic parameters refer to the observer's position or orientation, or to the zero-of-time on the observer's watch. There are seven extrinsic parameters: the four parameters t_0, μ_S, ϕ_S, and D correspond to the freedom to translate the same binary in

[3]In reality, radiation reaction will impose a small time variation in λ; however, this variation is known to be very small (See Ref. (**?**)) and we shall ignore it here. When a model of the time-variation of λ is eventually at hand, it would be trivial to generalize our treatment to incorporate it: one would just need an equation for $d\lambda/dt$, and in the parameter list λ would be replaced by λ_0—the value of λ at time t_0.

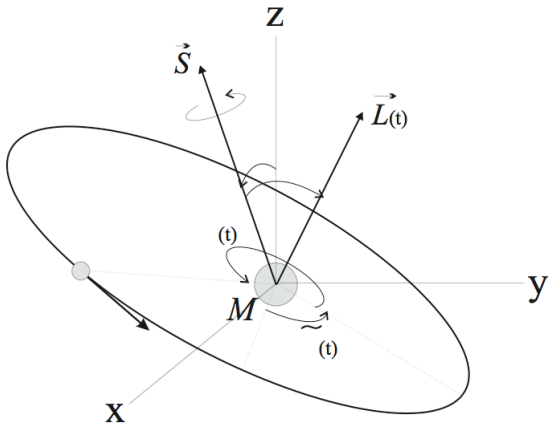

Figure 3.1: The MBH-CO system: setup and notation. M and μ are the masses of the MBH and the CO, respectively. The axes labeled $x{-}y{-}z$ represent a Cartesian system *based on ecliptic coordinates* (the Earth's motion around the Sun is in the x–y plane). The spin $\vec{S}$ of the MBH is parametrized by its magnitude S and the two angular coordinates θ_K, ϕ_K, defined (in the standard manner) based on the system $x{-}y{-}z$. $\vec{L}(t)$ represents the (time-varying) orbital angular momentum; its direction is parametrized by the (constant) angle λ between $\vec{L}$ and $\vec{S}$, and by an azimuthal angle $\alpha(t)$ (not shown in the figure). The angle $\tilde{\gamma}(t)$ is the (intrinsic) direction of pericenter, as measured with respect to $\vec{L} \times \vec{S}$. Finally, $\Phi(t)$ denotes the mean anomaly of the orbit, i.e., the average orbital phase with respect to the direction of pericenter.

space and time, and the three parameters μ_K, ϕ_K, and α_0 are basically Euler angles that specify the orientation of the orbit with respect to the observer (at t_0). The intrinsic parameters are the ones that control the detailed dynamical evolution of the system, without reference to the observer's location or orientation. In our parametrization, the seven intrinsic parameters are $\ln \mu$, $\ln M$, S/M^2, $\cos\lambda$, e_0, $\tilde{\gamma}_0$, and Φ_0. BCV observed (in the context of circular-orbit binaries with spin) that extrinsic parameters are generally much "cheaper" to search over than intrinsic parameters, which can be important for constructing efficient search strategies.

3.5.3 Principal axes

Let $\hat{n}$ be the unit vector pointing from the detector to the source, and let $\hat{L}(t)$ be the unit vector along the CO's orbital angular momentum. We find it convenient to work in a (time-varying) wave frame defined with respect to $\hat{n}$ and $\hat{L}(t)$. We define unit vectors $\hat{p}$ and $\hat{q}$ by

$$\begin{aligned} \hat{p} &\equiv (\hat{n} \times \hat{L})/|\hat{n} \times \hat{L}|, \\ \hat{q} &\equiv \hat{p} \times \hat{n}, \end{aligned} \tag{3.115}$$

we define two polarization basis tensors, just as in eq. (**??**):

$$\begin{aligned} e^{+}_{ij}(t) &\equiv \hat{p}_i \hat{p}_j - \hat{q}_i \hat{q}_j, \\ e^{\times}_{ij}(t) &\equiv \hat{p}_i \hat{q}_j + \hat{q}_i \hat{p}_j. \end{aligned} \tag{3.116}$$

The general GW strain field at the detector can then be written as

$$h_{ij}(t) = A^{+}(t) e^{+}_{ij}(t) + A^{\times}(t) e^{\times}_{ij}(t), \tag{3.117}$$

where $A^{+}(t)$ and $A^{\times}(t)$ are the amplitudes of the two polarizations. In the next section we derive expressions for $A^{+}(t)$ and $A^{\times}(t)$.

λ^0	t_0	t_0 is time where orbital frequency sweeps through fiducial value (e.g., 1mHz)
λ^1	$\ln\mu$	(ln of) CO's mass
λ^2	$\ln M$	(ln of) MBH's mass
λ^3	S/M^2	magnitude of (specific) spin angular momentum of MBH
λ^4	e_0	$e(t_0)$, where $e(t)$ is the orbital eccentricity
λ^5	$\tilde{\gamma}_0$	$\tilde{\gamma}(t_0)$, where $\tilde{\gamma}(t)$ is the angle (in orbital plane) between $\hat{L}\times\hat{S}$ and pericenter
λ^6	Φ_0	$\Phi(t_0)$, where $\Phi(t)$ is the mean anomaly
λ^7	$\mu_S\equiv\cos\theta_S$	(cosine of) the source direction's polar angle
λ^8	ϕ_S	azimuthal direction to source
λ^9	$\cos\lambda$	$\hat{L}\cdot\hat{S}(=\text{const})$
λ^{10}	α_0	$\alpha(t_0)$, where $\alpha(t)$ is the azimuthal direction of $\hat{L}$ (in the orbital plane)
λ^{11}	$\mu_K\equiv\cos\theta_K$	(cosine of) the polar angle of MBH's spin
λ^{12}	ϕ_K	azimuthal direction of MBH's spin
λ^{13}	$\ln(\mu/D)$	(ln of) CO's mass divided by distance to source

Table 3.2: Summary of physical parameters and their meaning. The angles (θ_S,ϕ_S) and (θ_K,ϕ_K) are associated with a spherical coordinate system attached to the ecliptic. $\hat{L}$ and $\hat{S}$ are unit vectors in the directions of the orbital angular momentum and the MBH's spin, respectively. For further details see figure 3.1 and the description in the text.

3.5.4 Peters-Matthews waveforms

In the quadrupole approximation, the metric perturbation far from the source is given (in the "transverse/traceless" gauge) by (35)

$$h_{ij}=(2/D)\big(P_{ik}P_{jl}-\frac{1}{2}P_{ij}P_{kl}\big)\ddot{I}^{kl} \tag{3.118}$$

where D is the distance to the source, the projection operator P_{ij} is given by $P_{ij}\equiv\eta_{ij}-\hat{n}_i\hat{n}_j$, and $\ddot{I}^{ij}$ is the second time derivative of the inertia tensor. In this paper we work in the limit of small mass ratio, $\mu/M\ll 1$, where μ and M are the masses of the CO and MBH, respectively. In this limit, the inertia tensor is just $I^{ij}(t)=\mu\, r^i(t)r^j(t)$, where $\vec{r}$ is the position vector of the CO with respect to the MBH.

Consider now a CO-MBH system described as a Newtonian binary, with semi-major axis a, eccentricity e, and orbital frequency $\nu=(2\pi M)^{-1}(M/a)^{3/2}$. Let $\hat{e}_1$ and $\hat{e}_2$ be orthonormal vectors pointing along the major and minor axes of the orbital ellipse, respectively. Since the orbit is planar, I^{ij} has only 3 independent components: I^{11}, I^{21}, and I^{22}, and as the motion is periodic, we can express I^{ij} as a sum of harmonics of the orbital frequency ν: $I^{ij}=\sum_n I^{ij}_n$.

We next denote

$$\begin{aligned} a_n &\equiv \frac{1}{2}(\ddot{I}^{11}_n-\ddot{I}^{22}_n),\\ b_n &\equiv \ddot{I}^{12}_n,\\ c_n &\equiv \frac{1}{2}(\ddot{I}^{11}_n+\ddot{I}^{22}_n). \end{aligned} \tag{3.119}$$

Peters and Matthews showed (31) that

$$\begin{aligned} a_n &= -n\mathcal{A}\big[J_{n-2}(ne)-2eJ_{n-1}(ne)+(2/n)J_n(ne)+2eJ_{n+1}(ne)-J_{n+2}(ne)\big]\cos[n\Phi(t)],\\ b_n &= -n\mathcal{A}(1-e^2)^{1/2}\big[J_{n-2}(ne)-2J_n(ne)+J_{n+2}(ne)\big]\sin[n\Phi(t)],\\ c_n &= 2\mathcal{A}J_n(ne)\cos[n\Phi(t)], \end{aligned} \tag{3.120}$$

where

$$\mathcal{A} \equiv (2\pi\nu M)^{2/3}\frac{\mu}{D}, \tag{3.121}$$

J_n are Bessel functions of the first kind, and $\Phi(t)$ is the mean anomaly (measured from pericenter). For a strictly Newtonian binary we have

$$\Phi(t) = 2\pi\nu(t - t_0) + \Phi_0, \tag{3.122}$$

where Φ_0 is the mean anomaly at t_0. Decomposing Eq. (3.117) into n-harmonic contributions and using Eq. (3.118), one then easily obtains explicit expressions for the n-harmonic components of the two polarization coefficients,

$$A^+ \equiv \sum_n A_n^+ \tag{3.123}$$

$$A^\times \equiv \sum_n A_n^\times\,. \tag{3.124}$$

$$\tag{3.125}$$

where the $A_n^{+,\times}$ are

$$\begin{aligned} A_n^+ &= -[1 + (\hat{L}\cdot\hat{n})^2]\,[a_n\cos(2\gamma) - b_n\sin(2\gamma)] + [1 - (\hat{L}\cdot\hat{n})^2]c_n, \\ A_n^\times &= 2(\hat{L}\cdot\hat{n})\,[b_n\cos(2\gamma) + a_n\sin(2\gamma)]\,, \end{aligned} \tag{3.126}$$

where γ is an azimuthal angle measuring the direction of pericenter with respect to $\hat{x} \equiv [-\hat{n} + \hat{L}(\hat{L}\cdot\hat{n})]/[1 - (\hat{L}\cdot\hat{n})^2]^{1/2}$. In practice, we truncate the sums in Eq. (3.123) at $n = 20$.

The angular momenum direction vector $\hat{L}$ is not constant, since $\hat{L}$ precesses about the MBH's spin direction $\hat{S}$. Let $\theta_L(t), \phi_L(t)$ be the angles specifying the instantaneous direction of $\hat{L}$. These can be expressed in terms of the other angles in the problem as follows. Recall that θ_K, ϕ_K give the direction of $\vec{S}$ in the ecliptic-based system ('K' standing for 'Kerr'). Let λ be the angle *between* $\hat{L}$ and $\hat{S}$, and let $\alpha(t)$ be an azimuthal angle (in the orbital plane) that measures the precession of $\hat{L}$ *around* $\hat{S}$: Specifically, let

$$\hat{L} = \hat{S}\,\cos\lambda + \frac{\hat{z} - \hat{S}\cos\theta_K}{\sin\theta_K}\sin\lambda\cos\alpha + \frac{\hat{S}\times\hat{z}}{\sin\theta_K}\,\sin\lambda\sin\alpha, \tag{3.127}$$

where $\hat{z}$ is a unit vector normal to the ecliptic. Then the angles $\theta_L(t), \phi_L(t)$ are given in terms of θ_K, ϕ_K, λ, $\alpha(t)$ by

$$\begin{aligned} \cos\theta_L(t) &= \cos\theta_K\cos\lambda + \sin\theta_K\sin\lambda\cos\alpha(t), \\ \sin\theta_L(t)\cos\phi_L(t) &= \sin\theta_K\cos\phi_K\cos\lambda - \cos\phi_K\cos\theta_K\sin\lambda\cos\alpha(t) + \sin\phi_K\sin\lambda\sin\alpha(t), \\ \sin\theta_L(t)\sin\phi_L(t) &= \sin\theta_K\sin\phi_K\cos\lambda - \sin\phi_K\cos\theta_K\sin\lambda\cos\alpha(t) - \cos\phi_K\sin\lambda\sin\alpha(t). \end{aligned} \tag{3.128}$$

The evolution equation for $\alpha(t)$ is given in Sec. 3.5.6 below.

3.5.5 The pericenter angle $\tilde{\gamma}$

As mentioned above, the angle γ that appears in Eqs. (3.126) measures the direction of pericenter with respect to $\hat{x} \equiv [-\hat{n} + \hat{L}(\hat{L}\cdot\hat{n})]/[1 - (\hat{L}\cdot\hat{n})^2]^{1/2}$. With this definition, γ is neither purely extrinsic nor purely intrinsic. (In the terminology of BCV, "intrinsic" parameters describe the system without reference to the location or orientation of the observer.) We will find it convenient

to introduce a somewhat different convention for the zero-point of this angle: We shall define $\tilde{\gamma}$ to be the direction of pericenter with respect to $\hat{L}\times\hat{S}$. Then $\tilde{\gamma}$ is a purely intrinsic quantity.

Clearly, γ and $\tilde{\gamma}$ are related by

$$\gamma = \tilde{\gamma} + \beta, \tag{3.129}$$

where β is the angle from $\hat{x}\propto[\hat{L}(\hat{L}\cdot\hat{n})-\hat{n}]$ to $(\hat{L}\times\hat{S})$. It is straightforward to show that β is given by

$$\begin{aligned}\sin\beta &= \frac{\cos\lambda\,\hat{L}\cdot\hat{n}-\hat{S}\cdot\hat{n}}{\sin\lambda\big[1-(\hat{L}\cdot\hat{n})^2\big]^{1/2}},\\ \cos\beta &= \frac{\hat{n}\cdot(\hat{S}\times\hat{L})}{\sin\lambda\big[1-(\hat{L}\cdot\hat{n})^2\big]^{1/2}}.\end{aligned} \tag{3.130}$$

To evaluate $\beta(t)$ in practice, it is useful to know the following relations:

$$\hat{S}\cdot\hat{n} = \cos\theta_S\cos\theta_K + \sin\theta_S\sin\theta_K\cos(\phi_S-\phi_K), \tag{3.131}$$

$$\hat{n}\cdot(\hat{S}\times\hat{L}) = \sin\theta_S\sin(\phi_K-\phi_S)\sin\lambda\cos\alpha + \frac{\hat{S}\cdot\hat{n}\cos\theta_K-\cos\theta_S}{\sin\theta_K}\sin\lambda\sin\alpha, \tag{3.132}$$

and

$$\hat{L}\cdot\hat{n} = \hat{S}\cdot\hat{n}\cos\lambda + \frac{\cos\theta_S-\hat{S}\cdot\hat{n}\cos\theta_K}{\sin\theta_K}\sin\lambda\cos\alpha + \frac{(\hat{S}\times\bar{z})\cdot\hat{n}}{\sin\theta_K}\sin\lambda\sin\alpha, \tag{3.133}$$

or, equivalently,

$$\hat{L}\cdot\hat{n} = \cos\theta_S\cos\theta_L + \sin\theta_S\sin\theta_L\cos(\phi_S-\phi_L). \tag{3.134}$$

Note that the time-variation of $\hat{S}\cdot\hat{n}$ is very small in the extreme mass-ratio case considered here. In our kludged model we approximate $\hat{S}$—and hence $\hat{S}\cdot\hat{n}$—as strictly constant.

3.5.6 Orbital evolution equations

We evolve $\Phi(t)$, $\nu(t)$, $\tilde{\gamma}(t)$, $e(t)$, and $\alpha(t)$ using the following PN formulae:

$$\frac{d\Phi}{dt} = 2\pi\nu, \tag{3.135}$$

$$\begin{aligned}\frac{d\nu}{dt} &= \frac{96}{10\pi}(\mu/M^3)(2\pi M\nu)^{11/3}(1-e^2)^{-9/2}\big\{\big[1+(73/24)e^2+(37/96)e^4\big](1-e^2)\\ &\quad+(2\pi M\nu)^{2/3}\big[(1273/336)-(2561/224)e^2-(3885/128)e^4-(13147/5376)e^6\big]\\ &\quad-(2\pi M\nu)(S/M^2)\cos\lambda(1-e^2)^{-1/2}\big[(73/12)+(1211/24)e^2\\ &\quad+(3143/96)e^4+(65/64)e^6\big]\big\},\end{aligned} \tag{3.136}$$

$$\begin{aligned}\frac{d\tilde{\gamma}}{dt} &= 6\pi\nu(2\pi\nu M)^{2/3}(1-e^2)^{-1}\left[1+\frac{1}{4}(2\pi\nu M)^{2/3}(1-e^2)^{-1}(26-15e^2)\right]\\ &\quad-12\pi\nu\cos\lambda(S/M^2)(2\pi M\nu)(1-e^2)^{-3/2},\end{aligned} \tag{3.137}$$

$$\begin{aligned}\frac{de}{dt} &= -\frac{e}{15}(\mu/M^2)(1-e^2)^{-7/2}(2\pi M\nu)^{8/3}\big[(304+121e^2)(1-e^2)\big(1+12(2\pi M\nu)^{2/3}\big)\\ &\quad-\frac{1}{56}(2\pi M\nu)^{2/3}\big((8)(16705)+(12)(9082)e^2-25211e^4\big)\big]\\ &\quad+e(\mu/M^2)(S/M^2)\cos\lambda\,(2\pi M\nu)^{11/3}(1-e^2)^{-4}\left[(1364/5)+(5032/15)e^2+(263/10)e^4\right],\end{aligned} \tag{3.138}$$

$$\frac{d\alpha}{dt} = 4\pi\nu(S/M^2)(2\pi M\nu)(1-e^2)^{-3/2}. \tag{3.139}$$

Equations (3.136), (3.137), and (3.138) are from Junker and Schäfer (37), except (i) the second line of Eq. (3.137) is from Brumberg (38) (cf. our Appendix A), and the last term in Eq. (3.136)—the term $\propto S/M^2$—is from Ryan (39). Eq. (3.139) is from Barker and O'Connell (40). The dissipative terms $d\nu/dt$ and de/dt are given accurately through 3.5PN order (i.e., one order higher than 2.5PN order, where radiation reaction first becomes manifest). The non-dissipative equations, for $d\tilde{\gamma}/dt$ and $d\alpha/dt$, are accurate through 2PN order.[4]

In solving the above time-evolution equations, the initial values (at time t_0) of Φ, ν, $\tilde{\gamma}$, e, and α are just the parameters Φ_0, ν_0, $\tilde{\gamma}_0$, e_0, and α_0.

We use our PN Eqs. (3.138) and (3.136) to evolve $e(t)$ and $\nu(t)$ forward in time, up to the point when the CO plunges over the top of the effective potential barrier. For a point particle in Schwarzschild, the plunge occurs at $a_{\rm min} = M(6+2e)(1-e^2)^{-1}$ (42), so we set

$$\nu_{\rm max} = (2\pi M)^{-1}[(1-e^2)/(6+2e)]^{3/2}\,, \tag{3.140}$$

so we cut off the integration when ν reaches this $\nu_{\rm max}$.

3.5.7 The polarization angle

Eq. (3.117) express the waveform in terms of polarization tensors $e_{ij}^{+,\times}(t)$ that are time-varying. However to generate the LISA responses with Synthetic LISA or LISA Simulator, it is useful to re-express Eq. (3.117) in terms of fixed polarization tensors. How do we do this? First note that at any instant, within the Synthetic LISA conventions, the polarization angle is given by

$$\psi_{SL} = arctan\Big(\frac{\sin\theta_L \sin(\phi_S-\phi_L)}{\cos\theta_S \sin\theta_L \cos(\phi_S-\phi_L) - \cos\theta_L \sin\theta_S}\Big)\,. \tag{3.141}$$

I.e., for this polarization angle, $h_+ = A^+$ and $h_\times = A^\times$. To use Synthetic LISA, though, we want to choose some fixed ψ– call it ψ_0. (Again, this just corresponds to fixing the basis of polarization tensors. One could just set ψ_0 equal to 0, but we'll allow it to be arbitrary here.) With respect to this new basis, $h_{+,\times}(t)$ are given by

$$\begin{aligned} h_+(t) &= A^+(t)cos(2\psi_0 - 2\psi_{SL}(t)) + A^\times(t)sin(2\psi_0 - 2\psi_{SL}(t)) \\ h_\times(t) &= A^\times(t)cos(2\psi_0 - 2\psi_{SL}(t)) - A^+(t)sin(2\psi_0 - 2\psi_{SL}(t)) \end{aligned} \tag{3.142}$$

3.5.8 Putting the pieces together

The algorithm for constructing our analytic kludge EMRI waveforms is then: Fix some fiducial frequency ν_0 and choose waveform parameters

$$(t_0,\, \ln\mu,\, \ln M,\, S/M^2,\, e_0,\, \tilde{\gamma}_0,\, \Phi_0,\, \cos\theta_S,\, \phi_S,\, \cos\lambda,\, \alpha_0, \cos\theta_K,\, \phi_K, D)\,. \tag{3.143}$$

Solve the ODEs (3.135)–(3.139) for $\Phi(t)$, $\nu(t)$, $\tilde{\gamma}(t)$, $e(t)$, $\alpha(t)$. Use $e(t)$ and $\nu(t)$ to calculate $a_n(t), b_n(t), c_n(t)$ in Eqs. (3.120). Calculate $\theta_L(t), \phi_L(t)$ using Eqs. (3.128), and then calculate $\gamma(t)$ from $\tilde{\gamma}(t)$ using Eqs. (3.129)–(3.134). Calculate the amplitude coefficients $A_n^{+,\times}$ using Eqs. (3.126) and (3.134). Calculate ψ_{SL} using Eq. (3.141). Then calculate $h_{+,\times}(t)$ (for the ψ_0 of your choice)

[4]In fact, the equations for $d\tilde{\gamma}/dt$ and $d\alpha/dt$ are missing terms proportional to $(S/M^2)^2$, which, according to usual "order counting" are classified as 2PN. However, this usual counting is misleading when the central object is a spinning BH: Because BHs are ultracompact, their spins are smaller than suggested by the usual counting, and the missing terms $\propto (S/M^2)^2$ have, in fact, the same magnitude as 3PN terms. Similarly, the terms $\propto (S/M^2)$ in Eqs. (3.137) and (3.139) can be viewed as effectively 1.5PN terms.

using Eqs. (3.142). Finally, *Synthetic LISA* and/or the *LISA Simulator* are used to create the LISA TDI responses $X(t)$, $Y(t)$, $Z(t)$ corresponding to $h_{+,\times}(t)$

Note that, in our kludge treatment, pericenter precession and Lense-Thirring precession have no effect on the a_n, b_n, c_n. The effect of these motions is simply to rotate the binary system, which modifies the amplitude harmonics (via Eqs. 3.126) and the also polarization angle $\psi_{SL}(t)$. The latter enters the time-dependence of $h_{+,\times}(t)$ via Eqs. (3.142).

3.5.9 Implementation

First of all we should mention that we have fixed the initial LISA's phase: $\bar{\phi}_0$ was taken to be zero.

The waveform described here was implemented as C++ class `AKWaveform` with the following methods:

- `AKWaveform(float spin, float mu, float MBHmass, float tfin, float timestep)`. This is constructor with parameters: `spin` is dimensional (reduced) spin of MBH (between 0 and 1), `mu` is the mass of CO in $M_\odot$, `MBHmass` is the mass of MBH in $M_\odot$, `tfin` is the duration time in sec (inital time is assumed to be zero), and `timestep` defines sampling interval (in sec.).

- `SetSourceLocation(float thS, float phS, float thK, float phK, float D)`. In this function user defines location of the source through the following parameters: `thS` this is θ_S, `phS` this is ϕ_S, `thK` corresponds to θ_K and `phK` to ϕ_K, the distance to the source is `D` and it is given in pc.

- `EstimateInitialParams(float elso, float nulso, float ein, float nuin)` This function estimates inital frequency `nuin` and initial eccentricity `ein` for values given at plunge `elso` by integrating simplified evolution equations backwards.

- `EvolveOrbit(float nu0, float eccen, float gamma0, float Phi0, float al0, float lam)`. This method is used to evolve the orbit with initial conditions specified by `nu0, eccen, gamma0, Phi0, al0` and for a given `lam` $= \lambda$

- `GetOrbitalEvolution(Matrix<float>& time, Matrix<float>& Phit, Matrix<float>& nut, Matrix<float>& gammat, Matrix<float>& et, Matrix<float>& alt)`. Using this function user can have a look at the result of integration of eqns. (3.135 - 3.139)

- `GetFinalOrbitalParams(float& t, float& eend, float& nuend)`. This method returns parameter at plunge: `t` the final time (sec), `eend` final eccentricity, `nuend` final orbital azimuthal frequency. User should compute orbital motion before calling this function.

- `GetWaveform(float ps0, Matrix<float>& time, Matrix<float>& hplus, Matrix<float>& hcross)`. Finally this function will fill up 1-d matrices (vectors) with waveforms with initial polarization angle specified by `ps0`$= \psi_0$.

Note that the orbital evolution is decoupled from the source location, so one can compute the waveform for many direction on the sky, for the same intrinsic parameters (multiple call to `SetSourceLocation` and `GetWaveform`, with only one run of `EvolveOrbit`).

3.5.10 Choice of parameters for Challenge-1 EMRI data sets

Since EMRI searches are expected to be quite challenging computationally, and since available computer power is expected to be considerably greater when LISA flies (2016 or later), the EMRI Challenge-I signal parameters are drawn from an artificially narrow range. The presumption is that probably any EMRI search strategy that is effective with this artificially narrow range could *also* work as a search over an astrophysically realistic parameter range, when more computer power becomes available. See table.... for the exact parameter ranges used for the EMRI Challenge-1 data sets.

Chapter 4

Data formats, software and simulators

The Mock LISA Data Challenge Task Force has developed a single file format to describe the challenges (i.e., the year-long time series of TDI observables) *and* all the ingredients that go into them. Thus, this file format can be used in all the phases of the pipelines (source parameter generation, GW strain computation, TDI-observable assembly) that lead to the final challenge files. Different information may go into files used in different ways, but the following is the general abstract structure of a MLDC file, with all allowed sections (we will describe the practical implementation of this structure later in this section).

4.1 Abstract file structure of the MLDC file format, v. 1.0

4.1.1 Prolog with file metadata

Describes author, generation date, generating software and its version, any other comments. More metadata fields will be added if it becomes clear that they would be useful. Users are also free to add their own metadata (e.g., for private tracking of simulation files) here, as well as in other sections; the standard input/output software will not complain, and can be easily modified to access these private data.

4.1.2 LISA Data section

This section describes a model of the LISA geometry to be used in the simulators. Since current Challenges employ the pseudo-LISA orbits of Sec. 2.3, this section contains the parameters (time offset, initial position and orientation) needed to fully specify those. These parameters are fixed in the Challenges and in the simulator pipelines,[1] so this section is merely descriptive.

4.1.3 Noise Data section

This section describes models of the LISA noises to be used in the simulators. Current Challenges employ pseudo-random noises generated within the simulators, so this section contains the parameters needed to reproduce those, such as power spectral densities, generation timesteps (which set the effective bandwidth of noises), time offsets, pseudo-random generators and seeds, and interpolator definitions. The naming of the LISA noises is given in Table 4.1. If a noise is defined repeatedly in this section, it is understood that the definitions are those of its additive components. If a noise is not defined, it is understood that it is set to zero (this is the case of laser phase noises in Challenge

[1] *Synthetic LISA* does have the capability of parsing a pseudo-LISA specification and honoring its parameters.

	spacecraft 1 (left bench, right bench)	s/c 2 (L,R)	s/c 3 (L,R)
laser noises	C1, C1s	C2, C2s	C3, C3s
proof-mass noises	pm1, pm1s	pm2, pm2s	pm3, pm3s
photodetector noises	pd3, pdm2	pd1, pdm3	pd2, pdm1

Table 4.1: Naming of standard LISA noises. Laser noises and proof-mass noises take the index of the bench where they sit (the "s" in right-bench noises denotes the "*" sometimes used in their symbolic representation). Photodetector noises take the index of the oriented LISA link that they terminate (and the "m" stands for "−").

observable	descriptor (equivalent strain)	descriptor (fractional frequency offset)
X, Y, Z	Xp, Yp, Zp	Xf, Yf, Zf

Table 4.2: TDI observable naming scheme. Challenge 1 includes only TDI-1.0/TDI-1.5 unequal-arm–Michelson observables. The *LISA Simulator* produces equivalent-strain data, whose rms spectrum can be converted into an rms spectrum of phase by multiplying by 10^10 m and dividing by the speed of light. *Synthetic LISA* produces fractional-frequency-fluctuation data, whose rms spectrum can be converted into an rms spectrum of phase by dividing by $2\pi f$.

1). Since the noise parameters are fixed in the Challenges and in the simulator pipelines,[2] this section is merely descriptive.

4.1.4 Source Data section

This section describes one or more of the standard waveforms used in the challenges. The description may be parametric (in terms of the standard parameters defined above in the sections on source classes), or the sources may be described by their positions in the sky[3] and by arrays of h_+ and $h_\times$ strains at the SSB (the specification of the arrays includes their offset, cadence, and length).

In the current Challenge pipelines, MLDC files containing parametric source descriptions are first created (by programs galactic_xml, BBH_xml, EMRI_xml); separate utilities (galactic, BBH_p, EMRI_p) are then used to create MLDC files containing "raw strain" descriptions of the same sources, which are fed to the simulators to produce the final challenge files.[4]

4.1.5 TDI Data section

This section contains time series of the TDI observables, and is the crucial part of MLDC Challenge files. The specification of TDI arrays includes their offset, cadence, and length, as well as their descriptors, given in Table 4.2. MLDC files containing TDI arrays are the final product of the Challenge pipelines.

[2] However, *Synthetic LISA* does have the capability of parsing noise descriptions and creating the corresponding pseudo-random objects, down to the specification of seeds and interpolator lengths.

[3] *Synthetic LISA* honors also a polarization parameter that represents a rotation of the h_+ and $h_\times$ axes; this parameter is always set to zero in the *LISA Simulator*, with the understanding that the definition of polarization is already folded into the source parameters.

[4] *Synthetic LISA* is capable of handling parametric source files directly, but this capability is not used in the pipelines. Note also that the *LISA Simulator* currently deals with MLDC source files that contain a single source description.

4.2 lisaXML implementation of the MLDC file format, v. 1.0

The MLDC file format is implemented using XML (the eXtensible Markup Language), an industry-standard cousin of the well-known HTML. XML documents are textual, and they consist of a hierarchy of *elements*, written with opening and closing *tags*, and containing textual data; elements may also have *attributes*. In the following example, `taskforce` and `member` are tags (their closing version prefixed by a slash); the former functions as a container, the second as a textual-data element, further specified by the attribute `italian`:

```
<taskforce>
  <member italian="yes">
    Alberto Vecchio
  </member>
  <member italian="no">
    Neil Cornish
  </member>
</taskforce>
```

Note that whitespace and newlines have no meaning within XML, except when they appear in other textual data.

lisaXML is an *application* of XML based on XSIL (the eXtensible Scientific Interchange Language), developed at Caltech for use in LIGO and in the Digital Sky (see `http://www.cacr.caltech.edu/SDA/xsil`). XSIL uses very few simple elements: `<XSIL Name="..." Type="...">` as a hierarchical container, `<Param Name="..." Unit="...">` to describe parameters (including their units), and `<Array Name="..." Type="...">` to specify arrays. In lisaXML, array data is stored externally in binary files, which are specified using the `<Stream>` element (which accommodates also the description of binary files as big-endian or little-endian, providing for translation between different CPU platforms). The advantage of this arrangement is that file metadata and physical parameters are easily parseable and editable by humans (or by the powerful XML libraries available for most programming languages), while binary arrays can be saved and loaded very efficiently.

Perhaps the easiest way to outline the correspondence between the abstract MLDC format and its lisaXML implementation is by showing a commented lisaXML file containing all main MLDC sections. (XML comments begin with `<!--` and end with `-->`. They are not to be confused with XSIL `Comment` elements; the former are fully transparent to XML libraries, while the latter are processed as any other element.)

```
<?xml version="1.0"?>
<!-- All XML files begin with this line. -->

<!DOCTYPE XSIL SYSTEM "http://www.vallis.org/lisa-xml.dtd">
<!-- This line specifies the "Document Type Definition", another XML file that
     describes the legal succession of XML tags in the document. The DTD file
     can be used to validate the syntax of all lisaXML files.
     This validation is only syntactic: it checks the well-formedness of elements,
     and the legal nesting of elements, but it cannot verify, for instance, that
     all parameters needed to specify a source are present. -->

<?xml-stylesheet type="text/xsl" href="lisa-xml.xsl"?>
```

```
<!-- This is the lisaXML XSL stylesheet, which allows standard-compliant web
     browsers such as Firefox and Safari to render the lisaXML file as nicely
     formatted HTML. Param sequences become tables, etc. -->

<XSIL>
    <!-- This is the main XSIL "container" for this file -->

    <!-- The Prolog appears at the beginning of the main XSIL container, but is
         not a container in itself -->

    <Param Name="Author">Michele Vallisneri</Param>


    <Param Name="GenerationDate" Type="ISO-8601">2006-03-10T14:53:34PST</Param>
    <!-- The ISO-8601 date is a standard with many options; it is used in
         lisaXML at the highest level of detail, except for sub-second timing.
         -->


    <Comment>
        This file produced by Synthetic LISA v. 1.3.1
    </Comment>

    <XSIL Type="LISAData">
        <!-- This XSIL container implements the LISA Data section. Currently
             only PseudoLISA specifications are supported. -->

        <XSIL Type="PseudoLISA">
            <Param Name="TimeOffset" Unit="Second">0.0</Param>
            <Param Name="InitialPosition" Unit="Radian">0.0</Param>
            <Param Name="InitialRotation" Unit="Radian">0.0</Param>
        </XSIL>


    <XSIL Type="NoiseData">

        <!-- This XSIL container implements the LISA Noise section. Currently
             only PseudoRandomNoise specifications are supported. -->

        <XSIL Name="pm1" Type="PseudoRandomNoise">
            <Param Name="SpectralType" Unit="String">WhiteAcceleration</Param>
            <!-- WhitePhase, WhiteFrequency, RedAcceleration (1/f) are also
                 supported -->

            <Param Name="Cadence" Unit="Second">100</Param>

            <Param Name="TimeOffset" Unit="Second">347.0</Param>
            <!-- Necessary to allow for TDI observables to be computable at
```

```
                time 0, since they involve delayed noises -->

            <Param Name="PowerSpectralDensity" Unit="(f/Hz)^n/Hz">2.5e-48</Param>

            <!-- The following supported by Synthetic LISA -->

            <Param Name="PseudoRandomGenerator" Unit="String">taus2-gsl1.4</Param>
            <Param Name="PseudoRandomSeed" Unit="1">224132777</Param>

            <Param Name="Interpolator" Unit="String">Lagrange</Param>
            <Param Name="InterpolatorWindow" Unit="1">4</Param>
</XSIL>

<!-- More PseudoRandomNoise sections would normally follow here. -->
   </XSIL>

   <XSIL Type="SourceData">
       <!-- This XSIL container implements the Source section. Currently
            only PlaneWave and SampledPlaneWave specifications are supported. -->

       <XSIL Name="AMCVn" Type="PlaneWave">
           <!-- An example of a parametric source definition -->

           <Param Name="SourceType" Unit="String">GalacticBinary</Param>

           <Param Name="EclipticLatitude" Unit="Radian">0.917311</Param>
           <Param Name="EclipticLongitude" Unit="Radian">2.97378</Param>
           <Param Name="Polarization" Unit="Radian">2.46531</Param>

           <Param Name="Amplitude" Unit="1">1.35695e-22</Param>
           <Param Name="Inclination" Unit="Radian">0.7679</Param>
           <Param Name="Frequency" Unit="Hertz">0.001944144722</Param>
           <Param Name="InitialPhase" Unit="Radian">0</Param>
       </XSIL>

       <XSIL Name="AMCVn" Type="SampledPlaneWave">
           <!-- An example of a "raw strain" source definition, corresponding
                to the AMCVn parametric source given above. Normally the
                two would not sit in the same lisaXML file. -->

           <Param Name="EclipticLatitude" Unit="Radian">0.917311</Param>
           <Param Name="EclipticLongitude" Unit="Radian">2.97378</Param>

           <Param Name="Polarization" Unit="Radian">0</Param>
           <!-- Since this applies only to Synthetic LISA, in the MLDC
                pipelines the polarization rotation is applied at the time
                of generating raw strains, and this parameter is therefore
                set to 0. -->
```

```
        <XSIL Name="hp,hc" Type="TimeSeries">
            <!-- All lisaXML strain and TDI arrays are specified using
                 the Array element, but they are enclosed in XSIL TimeSeries
                 containers that give details such as time offset and
                 duration. -->

            <Param Name="TimeOffset" Unit="Second">-900</Param>
            <!-- This -900 s time offset is always imposed in MLDC
                 raw strain files to allow for the formation of TDI
                 combinations, and for the retardation of the hp and hc
                 data (defined at the SSB) to the LISA s/c positions. -->

            <Param Name="Cadence" Unit="Second">15</Param>
            <Param Name="Duration" Unit="Second">31459080</Param>

            <Array Name="hp,hc" Type="double" Unit="1">
                <!-- The names of the quantities contained in the array
                     are given as a comma-separated string -->

                <Dim Name="Length">2097272</Dim>
                <Dim Name="Records">2</Dim>
                <!-- Thus the array consists of 2097272 rows, and
                      2 columns; the binary file ordering is row-first. -->

                <Stream Type="Remote" Encoding="Binary,BigEndian">
                    <!-- The location of the binary file is defined with
                         respect to the main XML file. -->

                    AMCVn.bin
                </Stream>
            </Array>
        </XSIL>

        <!-- The following supported by Synthetic LISA -->

        <Param Name="Interpolator" Unit="String">Lagrange</Param>
        <Param Name="InterpolatorWindow" Unit="1">2</Param>
    </XSIL>
</XSIL>

<XSIL Type="SourceData">
    <!-- This XSIL container implements the TDI data section. -->

    <XSIL Name="t,Xf,Yf,Zf" Type="TDIObservable">
        <Param Name="DataType">FractionalFrequency</Param>
        <!-- Synthetic LISA generates FractionalFrequency data, while the
             LISA Simulator generates Strain data. -->
```

```
            <XSIL Name="t,Xf,Yf,Zf" Type="TimeSeries">
                <!-- This TimeSeries is defined much as the TimeSeries in
                     SampledPlaneWave above. -->

                <Param Name="TimeOffset" Type="Second">0</Param>
                <Param Name="Cadence" Unit="Second">15</Param>
                <Param Name="Duration" Unit="Second">983040</Param>

                <Array Name="t,Xf,Yf,Zf" Type="double">
                    <Dim Name="Length">65536</Dim>
                    <Dim Name="Records">4</Dim>
                    <Stream Type="Remote" Encoding="Binary,BigEndian">
                        challenge-0.bin
                    </Stream>
                </Array>
            </XSIL>
        </XSIL>
    </XSIL>
</XSIL>
```

4.3 lisaXML input/output libraries

4.3.1 C/C++

The LISAtools subversion archive on SourceForge includes a directory `lisaXML/io-C`, which contains lisaXML input-output routines in C based on Aaron Voisine's `ezxml`. These routines are documented in the `readxml.h` and `lisaxml.h` header files, and they allow the input and output of TDI timeseries, as well as raw-strain source objects. They return `TimeSeries` (or `LISASource`) C structures that contain all relevant parameters (such as `Cadence`s), and that point to the TDI and strain arrays, stored as separate C arrays for each column. For instance, one could do

```
TimeSeries *ts;
ts = getTDIdata("challenge.xml");
printf("'%s': %d doubles (every %g seconds); ",ts->Name, ts->Length,ts->Cadence);
/* TDI data is in timeseries->Data[i]->data[], with i=0 for t, i=1 for Xf, ... */
```

All memory allocation is handled by the routines, and functions are provided to free memory on exit. See `lisaXML/io-C/readxmlmain.c` and the files in `lisaXML/C-examples` for practical examples of the usage of these routines. If you use them from C++, remember to `#include` the C headers within `extern "C" { }` blocks, to compile the libraries with your C compiler (e.g., `gcc`), and to link with your C++ compiler (e.g., `g++`).

4.3.2 Python

lisaXML input/output routines are built into Synthetic LISA. Once that package is loaded (`from synthlisa import *`), all the sections of a lisaXML can be easily turned into the corresponding Synthetic LISA objects. For instance,

```
inputXML = readXML("challenge.xml")
lisa = inputXML.getLISAGeometry()
```

would return a `LISA` object already usable in Synthetic LISA operations. This object would also define all the parameters given in the XML (as strings, in a list with their units). For instance, `lisa.InitialPosition` could return `['0.0','Radian']`.

Noises can be loaded with `inputXML.getLISANoise()`, which will return a list of Synthetic LISA `Noise` objects, each defining all XML parameters, as well as its `name`. The method `getTDINoise()` attempts to return an entire `TDInoise` object (with all 18 standard noises), ready to be used to compute TDI observables.

As for TDI time series, these are loaded with `inputXML.getTDITimeSeries()`; the result is a list of Python dictionaries (one for each TimeSeries within the lisaXML), which define keys such as `'Cadence'`, and `'Data'`. The latter contains the TDI observables as a `Numeric` array:

```
obs = inputXML.getTDITimeSeries()

print obs[0]['Cadence']
X = obs[0]['Data'][:,1]
```

lisaXML writing is achieved by creating the equivalent Synthetic LISA objects, opening a lisaXML file with

```
outputXML = lisaXML("output.xml")
```

and then calling `outputXML.LISAData(lisa)`, `outputXML.NoiseData(noise)` (which takes also a `TDInoise` composite object), `outputXML.SourceData(wave)`, and finally

```
outputXML.TDIData(data,length,cadence,description)
```

with `data` a `Numeric` array, and `description` a comma-separated string such as `'t,Xf,Yf,Zf'`.

4.3.3 MATLAB

While a MATLAB library to read lisaXML is still under development, it is easy to look up the name of the TDI data binary and the number of records in the XML (using a web browser, or just an editor), and then load the array with the MATLAB commands

```
readfile = fopen(filename,'r','n');
obs = fread(readfile,[4,inf],'double');
```

where `'n'` specifies that the data are to be read in native format (i.e., with the enddianness of the local platform); it could also be `'l'` or `'b'` to specify little- or big-endian explicitly. Also, 4 is the number of records in the array. The TDI observables then become available as `obs(1,:)`, `obs(2,:)`, etc.

4.4 The Galactic Binary Pipeline

Software for the generation of data streams containing galactic binary systems are available at:

`http://sourceforge.net/projects/lisatools`

in the directory `lisatools/MLDCpipelines/pipeline-gb`.

The scripts and codes presented in this pipeline were cut from the v2.1.1 distribution of `The LISA Simulator`. The pipeline will create time series files and lisaXML files which will be suitable for input into either `The LISA Simulator` or `Synthetic LISA`.

4.4.1 Pipeline input files

To run the MLDC binary pipeline, you must create two base files which will serve as the initial data for the pipeline: **gb.txt** and **sources.txt**.

gb.txt: This is a file which enumerates the parameters of the galactic binaries of interest. The columns in the file should be:

- ▷ Col 1: Binary identifier (name or other identifying tag for the binary)
- ▷ Col 2: ecliptic LATITUDE (in radians)
- ▷ Col 3: ecliptic LONGITUDE (in radians)
- ▷ Col 4: principle polarization angle (in radians)
- ▷ Col 5: binary amplitude
- ▷ Col 6: inclination angle (in radians)
- ▷ Col 7: GW frequency of the binary (in Hz)
- ▷ Col 8: orbital phase (in radians)

Several caveats to note with this file's preparation:

- Column 2 is *ecliptic latitude*, ranging from $+90^\circ$ at the north ecliptic pole to -90° at the ecliptic south pole.
- The amplitude in Column 5 can be computed from standard binary parameters such as masses m_1 and m_2, the orbital separation r between binary components[5], and the distance D to the binary as $\mathcal{A} = 2(G^2/c^4)m_1m_2/(r \cdot D)$.

sources.txt: This is a file which simply has a list of the binary names which should be processed from "gb.txt".

Col 1: Binary identifier (name or other identifying tag for the binary)

The column in this file is identical to column 1 in "gb.txt" if you want all the sources to be processed.

4.4.2 Pipeline codes

There are two codes included in this pipeline, along with two scripts to compile and run the pipeline[6]. To use the pipeline, you simply have to run the two scripts. The two scripts are:

- `gbPipelineCompile`: Compiles the two codes in preparation for processing the list of binaries.
- `gbPipelineRun`: Runs the two codes to process the list of binaries into barycenter waveforms and lisaXML files.

Note that if you are creating data streams which have altered cadences or observation times, those elements will need to be changed in the `LISAconstants.h` file and the pipeline will need to be recompiled. Similar changes will also need to be made in the simulator you choose to use for creating the LISA signal output.

[5] In some instances the orbital period P is known rather than the orbital separation r. In these instances, an application of Kepler's Third Law is in order: $GM_{tot} = (2\pi/P)^2 r^3$.

[6] A third script is included, but it is called from the second script and simply processes the list of binaries being churned through the pipeline.

4.4.3 Pipeline output

This pipeline produces h+ and hx for each source at the Solar Barycenter and saves the time-series in the Binary subdirectory. The xml wrapper for each time-series is written to the XML subdirectory. These files will need to be provided to the LISA simulation software to produce a final LISA signal output suitable for data analysis research.

4.5 The massive black hole binary inspiral pipeline

Software for the generation of data streams containing in-spirals from massive black hole binary systems are available at:

`http://sourceforge.net/projects/lisatools`

in the directory `lisatools/MLDCpipelines/pipeline-bbh`.

Waveform generation codes are available in `lisatools/MLDCpipelines/Waveforms`.

4.6 Simulators

The simulators used for the MLDC are available at:

- The LISA Simulator: `http://www.physics.montana.edu/lisa` – the challenge version is 2.1.1
- Synthetic LISA: `http://www.vallis.org/syntheticlisa` – the challenge version is 1.3.1

Chapter 5

Challenge 1

The goal of Challenge-1 is to enable the development of the necessary data analysis infrastructure of and building blocks for LISA data analysis. It focuses on the analysis of data sets that contain a single source or a small number of non-overlapping sources, although an exception is made for the data sets 1.1.4 and 1.1.5; see later. The first round of challenges concentrates on sources listed (currently) as minimum science requirements, *i.e.* galactic binaries, including verification binaries, and massive black hole binary systems. Data sets containing EMRIs will also be made available, in order to facilitate the development of data analysis tool for this particular demanding class of sources. However, the results of the challenge are due in June 2007 (the date for the round 2 of the MLDC) and not 1 December 2006.

For each of the challenge data sets a corresponding 'training data set will also be released the training data set is just a different instance of the mock data with a signal generated with the same properties (but different parameters) of that present in the (blind) challenge data set for which all the parameters are made public. The training data sets, as well as the challenge data sets are posted at:

`http://astrogravs.gsfc.nasa.gov/docs/mldc/round1.html`

The link "key" points to the xml file used for the signal generation and contain full information about the source(s).

In the following we provide details on each of the Challenge data sets for round 1.

5.1 Challenge 1.1 – Galactic binaries

In the following the notation $U[a, b]$ stands is short-hand for a uniform distribution over the range a, b.

C1.1.1a – Single source A 1-yr long data stream consisting of instrumental noise (Gaussian and stationary) according to the model reported in Section 2.6 and *one* low frequency galactic binary whose frequency is *exactly* monochromatic in the source reference frame. The SNR is in the range $\approx 10 - 20$.

Challenge 1.1.1a – Source(s): 1 galactic binary	
Parameters	Value or range
θ	with $\cos\theta$ chosen from U[-1,1]
φ	U$[0, 2\pi]$
ι	with $\cos\iota$ chosen from U[-1,1]
ψ	U$[0, 2\pi]$
$\mathcal{A}$	random, but such that SNR > 10
f_0	U[0.9 mHz, 1.1 mHz]
$\dot{f}_0$	0
$\ddot{f}_0$	0
ϕ_0	U$[0, 2\pi]$

C1.1.1b – Single source A 1-yr long data stream consisting of instrumental noise (Gaussian and stationary) according to the model reported in Section 2.6 and *one* mid-frequency galactic binary whose frequency is *exactly* monochromatic in the source reference frame. The SNR is in the range $\approx 10 - 20$.

Challenge 1.1.1b – Source(s): 1 galactic binary	
Parameters	Value or range
θ	with $\cos\theta$ chosen from U[-1,1]
φ	U$[0, 2\pi]$
ι	with $\cos\iota$ chosen from U[-1,1]
ψ	U$[0, 2\pi]$
$\mathcal{A}$	random, but such that SNR > 10
f_0	U[2.9 Hz, 3.1 Hz]
$\dot{f}_0$	0
$\ddot{f}_0$	0
ϕ_0	U$[0, 2\pi]$

C1.1.1c – Single source A 1-yr long data stream consisting of instrumental noise (Gaussian and stationary) according to the model reported in Section 2.6 and *one* high frequency galactic binary whose frequency is *exactly* monochromatic in the source reference frame. The SNR is in the range $\approx 10 - 20$.

Challenge 1.1.1c – Source(s): 1 galactic binary	
Parameters	Value or range
θ	with $\cos\theta$ chosen from U[-1,1]
φ	U$[0, 2\pi]$
ι	with $\cos\iota$ chosen from U[-1,1]
ψ	U$[0, 2\pi]$
$\mathcal{A}$	random, but such that SNR > 10
f_0	U[9 mHz, 11 mHz]
$\dot{f}_0$	0
$\ddot{f}_0$	0
ϕ_0	U$[0, 2\pi]$

Name	**ecl. lat. (rad)**	**ecl. long. (rad)**	f_{gw}(Hz) Training	f_{gw}(Hz) Blind Chal.
RXJ0806	0.2697997832	2.100953649	0.006220276603	0.005907385444
V407Vul	0.4353820782	5.147481878	0.00351250011	0.00367987089
ESCet	-0.1642146356	0.4258413339	0.003220611916	0.00332835297
AMCVn	0.6567450961	2.97249322	0.001944144722	0.00198757488
HPLib	-0.2481858196	4.101821899	0.001813236627	0.001777716317
EIPsc	0.1129174697	6.205956271	0.0005194805195	0.0005211327983
bltOPEN7	0.1996953268	4.757415	0.002095374315	0.002004086634
bltOPEN8	-0.6154346732	4.13981	0.001839178549	0.001767880534
bltOPEN9	-0.1422176732	4.649333	0.001910351046	0.001874145362
bltOPEN10	0.01476832679	4.756822	0.006334684838	0.006544361142
bltOPEN11	0.7497998268	5.174382	0.003208731343	0.003250185775
bltOPEN12	0.1200253268	4.780427	0.003502950886	0.00329556522
bltOPEN13	-0.7670126732	4.047208	0.002926858537	0.003062759803
bltOPEN14	-0.3583376732	4.46843	0.002706457364	0.002489978296
bltOPEN15	-0.2166256732	4.607265	0.002506720518	0.002355954312
bltOPEN16	-0.2676646732	4.50488	0.004841122821	0.004974343896
bltOPEN17	-0.06588667321	4.663938	0.003375130955	0.003190094085
bltOPEN18	0.04105132679	4.748964	0.004139974183	0.004146559486
bltOPEN19	-0.04352567321	4.749986	0.007870408994	0.008505086786
bltOPEN20	-0.01221367321	4.728495	0.002455096892	0.002351154304

Table 5.1: LISA Verification Binaries (Challenge 1.1.2. These data are the exact numbers used to produce the input data streams for the Challenge 1.1.2 (Verification Binaries). The two frequency columns indicate the slightly different values use in the training data set, and the blind challenge data. Note that of this list, only the first six binaries are known verification binaries at the current time. The values for their parameters are derived from a list maintained by Nelemans(**?**). The observer dependent parameters (ψ, ι, ϕ_o) for these six binaries are poorly known, and were assigned random values for this Challenge. The remaining binaries are included to mimic verification binaries which might be known before the LISA flight era; their parameter values are drawn from a population synthesis galaxy by Matt Benacquista.

C1.1.2 – Verification binaries We define "verification binaries" systems for which the GW frequency f_0 and source position in the sky θ and φ are exactly known. The data stream for this challenge is 1-yr long and contains instrumental noise (Gaussian and stationary) according to the model reported in Section 2.6 and GW signals from *20* verification binaries (6 "real" and 14 mock). The "name" and known parameters of each source are reported in Table 5.1. The signal of all the binaries is *exactly* monochromatic in the source reference frame ($\dot{f}_0 = \ddot{f}_0 = 0$). The remaining parameters are chosen randomly according to $\cos\iota \in$ U[-1,1], $\psi \in$ U$[0, 2\pi]$ and $\cos\iota \in$ U[-1,1]. The amplitude is selected randomly subject to the constraint that the SNR is > 10.

C1.1.3 – Resolvable binaries A 1-yr long data stream consisting of instrumental noise (Gaussian and stationary) according to the model reported in Section 2.6 and GW signals from *20* unknown resolvable galactic binaries. The signal of all the binaries is *exactly* monochromatic in the source reference. The SNR of each signal is > 10.

Challenge 1.1.3 – Source(s): 20 galactic binary	
Parameters	Value or range
θ	with $\cos\theta$ chosen from U[-1,1]
φ	U$[0, 2\pi]$
ι	with $\cos\iota$ chosen from U[-1,1]
ψ	U$[0, 2\pi]$
$\mathcal{A}$	random such that SNR > 10
f_0	U[0.1 mHz, 10.0 mHz]
$\dot{f}_0$	0
$\ddot{f}_0$	0
ϕ_0	U$[0, 2\pi]$

C1.1.4 – Overlapping binaries A 1-yr long data stream consisting of instrumental noise (Gaussian and stationary) according to the model reported in Section 2.6 and GW signals from ≈ 50 unknown overlapping galactic binaries. The sources located in a frequency region centered at 3 mHz over a 15μHz band. The signal of each binary is *exactly* monochromatic in the source reference. The SNR of each signal is > 5. Note that the total number of signals N is unknown, randomly chosen in the range $40 \leq N \leq 60$.

Challenge 1.1.4 – Source(s): ≈ 50 galactic binaries	
Parameters	Value or range
θ	with $\cos\theta$ chosen from U[-1,1]
φ	U$[0, 2\pi]$
ι	with $\cos\iota$ chosen from U[-1,1]
ψ	U$[0, 2\pi]$
$\mathcal{A}$	random such that SNR > 5
f_0	U[3.000 mHz, 3.015 mHz]
$\dot{f}_0$	0
$\ddot{f}_0$	0
ϕ_0	U$[0, 2\pi]$
number of sources	U[40,60]

C1.1.5 – Strongly Overlapping binaries A 1-yr long data stream consisting of instrumental noise (Gaussian and stationary) according to the model reported in Section 2.6 and GW signals from ≈ 40 unknown strongly overlapping galactic binaries. The sources located in a frequency region centered at 3 mHz over a $\pm 1.5\mu$Hz band. The signal of each binary is *exactly* monochromatic in the source reference. The SNR of each signal is > 5. Note that the total number of signals N is unknown, randomly chosen in the range $30 \leq N \leq 50$.

Challenge 1.1.5 – Source(s): ≈ 40 galactic binaries	
Parameters	Value or range
θ	with $\cos\theta$ chosen from U[-1,1]
φ	U$[0, 2\pi]$
ι	with $\cos\iota$ chosen from U[-1,1]
ψ	U$[0, 2\pi]$
$\mathcal{A}$	random such that SNR > 5
f_0	U[2.9985 mHz, 3.0015 mHz]
$\dot{f}_0$	0
$\ddot{f}_0$	0
ϕ_0	U$[0, 2\pi]$
number of sources	U[30,50]

5.2 Challenge 1.2 – Massive black hole binary systems

In the following the notation $U[a, b]$ stands is short-hand for a uniform distribution over the range a, b.

C1.2.1 – MBH binary inspiral A 1-yr long data stream consisting of instrumental noise (Gaussian and stationary) according to the model reported in Section 2.6 and *one* signal produced by a massive black hole binary inspiral – according to the model described in Section 3.4.6, restricted 2PN waveform – with coalescence taking place during the observation. The signal to noise ratio is ≈ 500.

Challenge 1.2.1 – Source(s): 1 MBH binary inspiral	
Parameters	Value or range
m_1	U[1,5] $\times 10^6\, M_\odot$
m_1/m_2	U[1,4]
θ	with $\cos\theta$ chosen from U[-1,1]
φ	U$[0, 2\pi]$
ι	with $\cos\iota$ chosen from U[-1,1]
ψ	U$[0, 2\pi]$
D	random, but such that SNR ≈ 500
ϕ_0	U$[0, 2\pi]$
t_c[1]	U[178 -20, 178 + 20] days

C1.2.2 – MBH binary inspiral A 1-yr long data stream consisting of instrumental noise (Gaussian and stationary) according to the model reported in Section 2.6 and *one* signal produced by a massive black hole binary inspiral – according to the model described in Section 3.4.6, restricted 2PN waveform – with coalescence taking place during the observation. The signal to noise ratio is $\approx 20 - 100$.

Challenge 1.2.2 – Source(s): 1 MBH binary inspiral	
Parameters	Value or range
m_1	U[1,5] $\times 10^6\, M_\odot$
m_1/m_2	U[1,4]
θ	with $\cos\theta$ chosen from U[-1,1]
φ	U$[0, 2\pi]$
ι	with $\cos\iota$ chosen from U[-1,1]
ψ	U$[0, 2\pi]$
D	random, but such that SNR $\approx 20 - 100$
ϕ_0	U$[0, 2\pi]$
${t_c}^2$	U[400 -40, 400 + 40] days

5.3 Challenge 1.3 – Extreme mass ratio inspirals

Coming very soon.....

Bibliography

[1] *LISA Simulator* v. 2.0, `www.physics.montana.edu/lisa`.

[2] N. J. Cornish and L. J. Rubbo, *Phys. Rev. D* **67**, 022001 (2003); erratum, *ibid.*, 029905 (2003). See also LISA Simulator v. 2.0, `www.physics.montana.edu/lisa`.

[3] L. J. Rubbo, N. J. Cornish, and O. Poujade, *Phys. Rev. D* **69**, 082003 (2004).

[4] *Synthetic LISA* v. 1.0, `www.vallis.org/syntheticlisa`.

[32] C. Cutler, *Phys. Rev. D* **57**, 7089 (1998).

[6] M. Vallisneri, the *TDI Rosetta Stone* v. 1/5/2005, `www.vallis.org/tdi`.

[7] M. Vallisneri, *Phys. Rev. D* **71**, 022001 (2005).

[8] J. W. Armstrong, F. B. Estabrook, and M. Tinto, *Astrophys. J.* **527**, 814 (1999); F. B. Estabrook, M. Tinto, and J. W. Armstrong, *Phys. Rev. D* **62**, 042002 (2000).

[9] N. J. Cornish & R. W. Hellings, *Clas. Quant. Grav.* **20**, 4851 (2003).

[10] D. A. Shaddock, M. Tinto, F. B. Estabrook, and J. W. Armstrong, *Phys. Rev. D* **68**, 061303(R) (2003); M. Tinto, F. B. Estabrook, and J. W. Armstrong, *Phys. Rev. D* **69**, 082001 (2004).

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

[12] A. Królak, M. Tinto, and M. Vallisneri, *Phys. Rev. D* **70**, 022003 (2004).

[13] Benacquista, M., DeGoes, J. & Lunder, D., *Class. Quant. Grav.* **21**, S509 (2004)

[14] Edlund, J. A., Tinto, M., Krolak, A. & Nelemans, G., *Phys. Rev. D* **71**, 122003 (2005)

[15] Hils, D., Bender, P. & Webbink R. F., *Astrophys. J.*, *Astrophys. J.* **360**, 75 (1990).

[16] Peters, P. C. & Mathews, J., *Phys. Rev.* **131**, 434 (1963).

[17] Pierro, V., Pinto, I. M., Spallicci, A. D., Laserra, E. & Recano, F., *Mon. Not. R. Astron. Soc.* **325**, 358 (2001)

[18] Timpano, S., Rubbo, L. & Cornish, N., *Preprint*, gr-qc/0504071 (2005).

[19] Buonanno, Chen, Vallisneri, Phys.Rev. D67 (2003) 024016

[20] Kidder, Phys.Rev. D52 (1995) 821

[21] Finn, Chernoff, Phys.Rev. D47 (1993) 2198

[22] Buonanno, Chen, Vallisneri, Phys.Rev D67 (2003) 104025

[23] Buonanno, Chen, Pan, Vallisneri Phys.Rev. D70 (2004) 104003

[24] Blanchet, Faye, Iyer, Joguet Phys.Rev. D65 (2002) 061501

[25] Damour, Iyer, Sathyaprakash Phys.Rev. D63 (2001) 044023

[26] Cutler, Phys.Rev. D57 (1998) 7089

[27] Apostolatos, Cutler, Sussman, Thorne, Phys.Rev D49 (1994) 6274

[28] Krolak, Tinto, Vallisneri, Phys.Rev D70, (2004) 022003

[29] L. Barack and C. Cutler, Phys. Rev. **69**, 082005(2004).

[30] F. D. Ryan, Phys. Rev. D **56**, 1845 (1997).

[31] P. C. Peters and J. Mathews, Phys. Rev. **131**, 435 (1963); P. C. Peters, Phys. Rev. **136**, B1224 (1964).

[32] C. Cutler, Phys. Rev. D. **57**, 7089 (1998).

[33] T. Apostolatos, C. Cutler, G. J. Sussman, and K. S. Thorne, Phys. Rev. D **49** 6274 (1994).

[34] A. Buonanno, Y. Chen, and M. Vallisneri, Phys. Rev. D **67**, 024016 (2003).

[35] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (Freeman, San Francisco, 1973), chapter 33.

[36] Marković, D. M. 1993, Phys. Rev. D 48, 4738.

[37] W. Junker and G. Schäfer, Mon. Not. R. astr. Soc. **254**, 146 (1992).

[38] V. A. Brumberg, *Essential Relativistic Celestial Mechanics* (IOP Publishing, Bristol, 1991).

[39] F. D. Ryan, Phys. Rev. D **53**, 3064 (1996).

[40] B. M. Barker and R. F. O'Connell, Phys. Rev. D **12**, 329 (1975).

[41] L. Blanchet, G. Faye, B. R. Iyer, and B. Joguet, Phys. Rev. D **65**, 061501 (2002).

[42] C. Cutler, D. Kennefick and E. Poisson, Phys. Rev. D **50**, 3816 (1994).

[43] C. Cutler, and E. E. Flanagan, Phys. Rev. D **49** 2658 (1994).

[44] L. E. Kidder, Phys. Rev. D **52**, 821 (1995).