

Massive Black Hole Binary Inspirals: Results from the LISA Parameter Estimation Taskforce

K. G. Arun^{1,2}, Stas Babak³, Emanuele Berti^{4,5}, Neil Cornish⁶, Curt Cutler^{4,7}, Jonathan Gair⁸, Scott A. Hughes⁹, Bala R. Iyer¹⁰, Ryan N. Lang⁹, Ilya Mandel¹¹, Edward K. Porter^{3,12}, Bangalore S. Sathyaprakash¹³, Siddhartha Sinha^{10,14}, Alicia M. Sintes^{3,15}, Miquel Trias¹⁵, Chris Van Den Broeck¹³, Marta Volonteri¹⁶

¹ LAL, Univ. Paris-Sud, IN2P3/CNRS, Orsay, France

² Institut d'Astrophysique de Paris, UMR 7095-CNRS, Université Pierre et Marie Curie, 98bis boulevard Arago, 75014 Paris, France

³ Max-Planck-Institut für Gravitationsphysik (Albert-Einstein-Institut), Am Mühlenberg 1, D-14476 Golm bei Potsdam, Germany

⁴ Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA

⁵ Dept. of Physics and Astronomy, The University of Mississippi, University, MS 38677-1848, USA

⁶ Dept. of Physics, Montana State Univ., Bozeman, MT 59717, USA

⁷ Theoretical Astrophysics, California Inst. of Technology, Pasadena, CA 91125

⁸ Institute of Astronomy, University of Cambridge, Cambridge, CB30HA, UK

⁹ Dept. of Physics and Kavli Institute for Astrophysics and Space Research, MIT, Cambridge, MA 02139

¹⁰ Raman Research Institute, Bangalore, 560 080, India

¹¹ Dept. of Physics and Astronomy, Northwestern Univ., Evanston, IL, USA

¹² APC (AstroParticules et Cosmologie), 10, rue Alice Domon et Léonie Duquet, 75205 Paris Cedex 13, France

¹³ School of Physics and Astronomy, Cardiff University, 5, The Parade, Cardiff, UK, CF24 3YB

¹⁴ Dept. of Physics, Indian Institute of Science, Bangalore, 560 012, India

¹⁵ Departament de Física, Universitat de les Illes Balears, Cra. Valldemossa Km. 7.5, E-07122 Palma de Mallorca, Spain

¹⁶ Dept. of Astronomy, University of Michigan, Ann Arbor, MI 48109

Abstract. The LISA Parameter Estimation (LISAPE) Taskforce was formed in September 2007 to provide the LISA Project with vetted codes, source distribution models, and results related to parameter estimation. The Taskforce's goal is to be able to quickly calculate the impact of any mission design changes on LISA's science capabilities, based on reasonable estimates of the distribution of astrophysical sources in the universe. This paper describes our Taskforce's work on massive black-hole binaries (MBHBs). Given present uncertainties in the formation history of MBHBs, we adopt four different population models, based on (i) whether the initial black-hole seeds are small or large, and (ii) whether accretion is efficient or inefficient at spinning up the holes. We compare four largely independent codes for calculating LISA's parameter-estimation capabilities. All codes are based on the Fisher-matrix approximation, but in the past they used somewhat different signal models, source parametrizations and noise curves. We show that once these differences are removed, the four codes give results in extremely close agreement with each other. Using a code that includes both spin precession and higher harmonics in the gravitational-wave signal, we carry out Monte Carlo simulations and determine the number of events that can be detected and accurately localized in our four population models.

1. Introduction

These proceedings report results on merging massive black-hole binaries (MBHBs) obtained by the LISA Parameter Estimation (LISAPE) Taskforce. The LISAPE Taskforce was established at the September, 2007, LISA International Science Team (LIST) meeting at ESTEC, under the auspices of the LIST’s Working Group 1b (Data Analysis). The LISAPE Taskforce was charged with developing a set of vetted tools for quickly estimating LISA’s science reach for various mission configurations and gravitational-wave (GW) sources.

The initial impetus for creating our Taskforce arose because several research groups had independently written codes to calculate LISA’s capabilities to extract the parameters of MBHBs [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14], but these groups’ published results appeared on their face to be discrepant. However, the various groups also used somewhat different approximations in their signal models, as well as somewhat different assumptions for the LISA noise curve (unfortunately, there has never been an “official” LISA noise curve). Therefore it was unclear whether the differing results simply reflected these different assumptions, or whether they were due to bugs in one or more codes.

The Taskforce’s first goal was to resolve that question. As described in Sec. 3, it turned out that there were no bugs: the differing results were primarily due to using different noise curves, and especially different low-frequency cut-offs. The second goal was to arrive at vetted parameter-estimation codes that the LISA Project could safely use in its work, e.g., in helping set LISA’s sensitivity requirements, as set forth in the evolving LISA Science Requirements Document [15].

In addition, the Taskforce wanted to develop a set of models for the distribution of MBHB events in the universe: event rates as a function of the black-hole masses, spins, and redshift, based on well-motivated assumptions regarding the birth and growth history of massive black holes in the universe. The goal was to generate realistic source distributions, which we could “feed” to our vetted parameter-estimation codes in Monte Carlo fashion, arriving at realistic ensembles of LISA observations and associated parameter-estimation accuracies. Our hope is that in the future other LISA researchers will use these same MBHB ensembles when evaluating LISA science performance, so that different researchers will be comparing “apples to apples”.

Of course, today our ignorance concerning MBHB birth and growth is rather humbling [16, 17, 18]. It therefore behooves us to consider a range of plausible assumptions. In the end, the Taskforce settled on four representative source distributions, constructed from four merger tree models. The four merger-tree models arise from two choices for astrophysical inputs: (1) whether the masses of the initial “seed” black holes are small or large, and (2) whether accretion is efficient or inefficient at spinning up the massive black holes. Many more models could of course be developed, but the Taskforce felt that these four do very broadly sample the range of possibilities, and so give useful insight into how much LISA’s science reach depends on the distribution of MBHBs in the universe.

The organization and procedures of the LISAPE taskforce were modelled on those of the quite successful Mock LISA Data Challenge (MLDC) Taskforce (see, e.g., [19]). Membership is open, we have group telecons roughly every two weeks, and the telecon minutes and other Taskforce documents are posted on our wiki:

<http://www.tapir.caltech.edu/dokuwiki/lisape:home>

It turned out that the creation of our Taskforce was quite timely, since by May, 2008, we were in a position to contribute significantly to a “descope exercise” that the LISA project carried out in that month at NASA’s behest. The results of that exercise are summarized in R. Stebbins’s contribution to this volume.

The organization of this paper is as follows. In Sec. 2 we describe the four merger-tree models that we use, and the underlying astrophysical assumptions. In Sec. 3 we describe the inter-code checks that we did to validate our parameter-estimation codes. All these codes calculate expected parameter uncertainties via the Fisher-matrix formalism [20]. Such results are accurate only up to corrections of order $1/\rho$, where ρ is the signal-to-noise ratio (SNR). However they are much faster to implement than more accurate Markov Chain Monte Carlo methods, and so more suitable for exploration of a large parameter space. For our results, presented in Sec. 4, we used the parameter-estimation code developed by the Montana/MIT group, since their model waveforms already incorporated both higher harmonics and spin-orbital precession effects. In Sec. 5 we summarize our conclusions and discuss briefly the planned future work of the LISAPE Taskforce.

2. Massive black-hole merger trees

The cosmological evolution of massive black holes can be determined by following the merger history of dark-matter halos and of the associated black holes by cosmological Monte Carlo realizations of the merger hierarchy from early times until the present in a Λ CDM cosmology ($H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$). The simulations provide the masses, redshifts and spins of merging black holes.

Two important sources of uncertainty in merger-tree models of black-hole formation are (i) the formation mechanism and mass of the first “seed” black holes, and (ii) the details of how accretion causes black holes to grow in time (see e.g. [16, 17, 18]). To bracket these uncertainties we focused on four representative models.

As a representative model with “light” black-hole seeds we considered the Volonteri-Haardt-Madau (VHM) scenario [21], where light seed black holes of $m_{\text{seed}} \sim$ few hundred M_\odot are produced as remnants of metal-free stars at redshift $z \gtrsim 20$. In an alternative scenario, “heavy” seeds with $m_{\text{seed}} \sim 10^5 M_\odot$ are formed as the end-product of dynamical instabilities arising in massive gaseous protogalactic disks in the redshift range $10 \lesssim z \lesssim 15$. To allow for the possibility of heavy seeds, we considered a variant of this scenario proposed by Begelman, Volonteri and Rees ([22], henceforth BVR). Both models (light and heavy seeds) can reproduce the AGN optical luminosity function in the redshift range $1 \lesssim z \lesssim 6$, but they result in different coalescence rates of MBHBs and hence in different GW backgrounds [17].

To bracket uncertainties in the growth of black holes by accretion, in both the “light seed” (VHM) and “heavy seed” (BVR) cases we considered two different accretion models. We adopted either the standard “prolonged accretion” scenario, where accretion of material with constant angular momentum axis rapidly spins up the holes [23, 24], or a “chaotic accretion” scenario [25], where accretion always proceeds via very small and short episodes, caused by fragmentation of the accretion disc where it becomes self-gravitating. Since counter-rotating material spins black holes down more efficiently than co-rotating material spins them up, and it is quite unlikely for mergers to produce rapidly spinning holes [26], the chaotic accretion scenario implies that black-hole spins are typically rather small [27].

For all four models (with heavy/light seeds and efficient/chaotic accretion), the

spin resulting from individual merger events was determined using a semianalytical fitting formula based on numerical relativity simulations of the merger process [28]. This formula provides the modulus of the final spin $|\mathbf{j}_{\text{fin}}|$ given the binary's mass ratio $q = M_1/M_2$ and the initial spins $\hat{\mathbf{j}}_i$ ($i = 1, 2$): $|\mathbf{j}_{\text{fin}}|(q, |\mathbf{j}_1|, |\mathbf{j}_2|, \cos\alpha, \cos\beta, \cos\gamma)$. The three angles $\cos\alpha = \hat{\mathbf{j}}_1 \cdot \hat{\mathbf{j}}_2$, $\cos\beta = \hat{\mathbf{j}}_1 \cdot \hat{\mathbf{L}}$ and $\cos\gamma = \hat{\mathbf{j}}_2 \cdot \hat{\mathbf{L}}$, where a hat denotes unit vectors, describe the orientation of the initial spins relative to the binary's orbital angular momentum \mathbf{L} (see [28] for details). The merger-tree simulations do not provide information on the source position in the sky and on the orientation of the binary's angular momentum. However, by averaging over different merger trees we can reasonably assume that all angles are isotropically distributed in the sky.

We further assume that spin alignment is not efficient, so that the spin directions at merger are isotropically distributed. This assumption may be violated if mergers commonly occur in gas-rich environments, and if the torque exerted by the gas is efficient in producing alignment of the black hole's angular momenta, as suggested in [29] (see [27] for more details).

2.1. Details of the merger-tree implementation

A merger tree traces the merger history that leads to a $z = 0$ dark-matter halo in a hierarchical cosmology. For each formation scenario, our sample is based on $N_{\text{tree}} \sim 10$ different halo masses [30] that span the masses of interest (from a dwarf galaxy to a cluster of galaxies). Each merger tree is characterized by the mass of the parent halo at $z = 0$ and its Press-Schechter weight $W_{\text{PS}}^{(k)}$ ($k = 1, \dots, N_{\text{tree}}$) [31, 32], which is used to scale the results to the (comoving) number density of sources. Furthermore, for each halo mass a different number $N_{\text{real}}^{(k)}$ of realizations of its merger history have been performed, to take into account cosmic variance. Typically, large-mass halos have a smaller Press-Schechter weight (inherent in the adopted cosmological model) and a smaller number of realizations (due to computational burden).

For each tree k we produced data files listing black-hole masses, spins and redshifts corresponding to “branches” of the tree where a merger event occurs. All quantities in these files are measured in the source frame, at variance with the convention used in the MLDC (recall that $M = (1 + z)M_{\text{source}}$). Including all merger trees and all realizations of each merger tree, in a typical model such as VHM we have at least $\sim 5 \times 10^4$ merging events, but many of these events will *not* be detectable by LISA. Once we choose a criterion to select detectable binaries, e.g. by requiring the SNR to be larger than some threshold value $\rho > \rho_{\text{th}}$ in a one-year observation time, the number of events at a given redshift z per comoving volume is

$$N_{\text{com}}(z) = \sum_{k=1}^{N_{\text{tree}}} \sum_{j=1}^{N_{\text{mergers}}^{(k)}} \frac{H(\rho^{(j,k)}(z) - \rho_{\text{th}}) \times W_{\text{PS}}^{(k)}}{N_{\text{real}}^{(k)}}, \quad (1)$$

where $H(x)$ is the Heaviside step function, and in practice we choose $\rho_{\text{th}} = 10$. The number of observable events at $z = 0$ per unit time and redshift is then given by

$$\frac{d^2N}{dzdt} = 4\pi c N_{\text{com}}(z) \left[\frac{D_L(z)}{(1+z)} \right]^2, \quad (2)$$

where D_L is the luminosity distance [33].

3. Code Comparisons for Massive Black Holes with Higher-Harmonic Corrections

The importance to GW astronomy of higher-harmonic corrections to post-Newtonian (PN) waveforms for massive black-hole inspirals has been recognized long ago [3, 4], but recently there have been several papers trying to quantify this effect more precisely [9, 10, 11, 13, 14]. Though the qualitative conclusions of all these papers were similar, the quantitative estimates were different. This has to do with differences in the nature of waveforms used in the analysis (time versus frequency domain), the different noise power spectral densities (PSDs) employed and different ways of truncating the PN waveforms. The main goal of the exercise presented in this section is to bring together the results of various groups and compare them using common values for the masses, the duration of the signal in the LISA band, the noise PSD and the lower frequency cut-off of the detector.

To this end it was decided that each group would analyze four different non-spinning MBHB sources. The sources were described by the following parameters: individual masses (m_1, m_2), co-latitude and longitude (θ_S, ϕ_S), the polarization and inclination angles (ψ, ι) or the orbital angular momentum variables (θ_L, ϕ_L), the luminosity distance D_L , the orbital phase at coalescence φ_c and the instant of coalescence t_c . The given parameter sets were $(m_1/M_\odot, m_2/M_\odot, \theta_S, \phi_S, \psi, \iota, \theta_L, \phi_L, \varphi_c, D_L/\text{Gpc}) = (3 \times 10^6, 10^6, 0.39845, 5.10773, 2.41199, 2.77508, 2.20726, 2.85098, 4.04657, 25.8416)$ for case 1, and $(3 \times 10^5, 10^5, 1.62168, 0.920401, 1.608812, 0.949798, 2.570387, 0.977606, 2.697954, 25.8416)$ for case 2. Note that all angular variables are given in radians, and that the chosen luminosity distance corresponds to a redshift $z = 3$. It was decided that each group would use waveforms which were correct to 2PN order in both amplitude and phase, and for each source we assumed one year of data. The time of coalescence is $t_c = 1$ year in cases 1a and 2a; in cases 1b and 2b, $t_c = 1.05$ years, so that the merger occurs after the end of the data set. We assumed common MLDC [41] conventions for the angles, and we adopted a common list of physical constants.

More importantly, all groups agreed to use the noise curve that was used in the second round of the MLDC [34]. There are two components to the noise model: instrumental noise (S_n) and confusion noise from the galactic foreground (S_{conf}). For our comparison exercise, the (sky-averaged) instrumental noise is

$$S_n(f) = \frac{1}{L^2} \left\{ \left[1 + \frac{1}{2} \left(\frac{f}{f_*} \right)^2 \right] S_p + \left[1 + \left(\frac{10^{-4}}{f} \right)^2 \right] \frac{4S_a}{(2\pi f)^4} \right\}, \quad (3)$$

where f is in Hz, $L = 5 \times 10^9$ m is the armlength, $S_p = 4 \times 10^{-22}$ m² Hz⁻¹ is the (white) position noise level, $S_a = 9 \times 10^{-30}$ m² s⁻⁴ Hz⁻¹ is the white acceleration noise level (assumed equal to the red acceleration noise level) and $f_* = c/(2\pi L)$ is the LISA arm transfer frequency. Note that the term $(f/f_*)^2/2$ in Eq. (3) is not strictly due to an increase in shot noise at high-frequency. Rather LISA's response to GWs falls rapidly for $f > f_*$, and in the low-frequency approximation adopted in our parameter estimation codes, this effect is typically accounted for by adjusting the noise curve in this fashion.

The galactic confusion noise is estimated by direct simulation. A Nelemans *et al.* [35, 36] population synthesis code was used to produce a catalog of galactic binaries with periods shorter than 2×10^4 seconds. The catalog contained 2.6×10^7 detached

binaries and 3.4×10^6 interacting binaries. The LISA response to this foreground was computed [37] and added to simulated instrument noise. A Bayesian model selection technique was then used to determine which systems could be individually identified and regressed from the data [37]. The residual was then smoothed using a Savitzky-Golay smoothing filter to produce an estimate for the galactic confusion noise. The confusion noise used in the comparisons is given by a piecewise fit to the residual for a two-year nominal lifetime without interacting binaries:

$$S_{\text{conf}}(f) = \begin{cases} 10^{-44.62} f^{-2/3} & (f \leq 10^{-3}), \\ 10^{-50.92} f^{-4.4} & (10^{-3} < f < 10^{-2.7}), \\ 10^{-62.8} f^{-8.8} & (10^{-2.7} < f < 10^{-2.4}), \\ 10^{-89.68} f^{-20} & (10^{-2.4} < f < 10^{-2}), \\ 0 & (f \geq 10^{-2}), \end{cases} \quad (4)$$

where f is in Hz. While there were differences between the way waveforms were treated, the main difference between the various groups was the analytical noise curve that was used. Once a common noise curve was agreed upon, most of the disagreement between groups disappeared. A final tidying up of the slight disagreements between waveform models produced a concrete verification of the results.

In the next few sections, we outline the different waveform approximations that were used. Two groups used time-domain waveforms, and two groups used frequency-domain waveforms. To begin with, it seemed reasonable to make comparisons between either time-domain or frequency-domain based codes. We then compared between the time- and frequency-domain results. The main differences between waveform models are summarized in Table 1.

Table 1. Comparisons of the different methods employed in the analysis. All waveforms used a 2 PN phase approximation for the study.

Group	AISSV	MM	CP	TS
Domain :	Frequency	Time	Time	Frequency
Amplitude:	Full PN expansion and truncation.	MLDC taper, Hann window.	MLDC taper.	Full PN expansion. No truncation.
Spin :	No	Yes	No	Yes
Spin Precession :	No	Yes	No	No

3.1. Time-Domain Waveforms

Two independent time-domain codes were used for the analysis. The first was used by Cornish and Porter (CP) [13] and is based on a 2PN waveform in both phase and amplitude, derived by Blanchet, Iyer, Will and Wiseman [38]. This code uses analytic expressions for the evolution of both the orbital phase and frequency. The code developed by the Montana/MIT group (MM) [39] is based on comparable-mass Kerr inspiral waveforms as developed by Apostolatos, Cutler, Sussman and Thorne [40]. This code used coupled ordinary differential equations to evolve the spin and spin-precession equations. For this study, the spin is set to zero. In order to dampen ringing in the Fourier domain, both waveforms use the standard MLDC hyperbolic truncation function [41]. This function begins to truncate the waveform at an orbital

separation of $7M$. The MM code also uses a Hann window to prevent leaking of higher harmonics into each other in the Fourier transform.

Table 2. Fractional differences in the claimed parameter-estimation accuracy between the CP and MM models. We list the SNR and the accuracy in (from left to right): individual masses, time to coalescence, luminosity distance, sky angles, orbital inclination, polarization angle and orbital phase at coalescence. Note that for each quantity $\Delta\lambda$, the value in the Table should be multiplied by 10^{-4} .

Case	$\Delta(\text{SNR})$	$\Delta \ln(m_1)$	$\Delta \ln(m_2)$	$\Delta \ln(t_c)$	$\Delta \ln(D_L)$	$\Delta \cos \theta_S$	$\Delta \phi_S$	$\Delta \iota$	$\Delta \psi$	$\Delta \varphi_c$
1a	1.25	5.61	5.84	6.27	175.07	14.57	30.09	248.35	125.69	57.29
1b	1.73	2.90	2.89	3.55	32.75	0.48	3.61	41.29	37.28	3.18
2a	18.49	19.99	20.05	74.95	566.11	365.62	90.57	800.23	385.82	146.92
2b	0.77	0.97	0.98	0.58	3.25	1.04	1.08	1.95	8.55	0.79

In Table 2 we show the difference in the claimed parameter-estimation accuracies between the two methods. We can see that this difference is typically of the order of a percent, and sometimes much less than that: for example, in case 1a the fractional difference between codes in calculated mass accuracy is $\sim 6 \times 10^{-4}$.

3.2. Frequency-Domain Waveforms

The two groups using frequency domain codes were Arun, Iyer, Sathyaprakash, Sinha and Van Den Broeck (AISSV) [9, 10] and Trias and Sintes (TS) [11, 14]. While both groups were using the Stationary Phase Approximation (SPA) [42] to obtain frequency-domain waveforms, the treatment of some aspects of waveform generation was different. The SPA amplitude of the k th harmonic is composed of a ratio of PN series. The AISSV group expanded the ratio of these series and truncated at the required 2 PN order. They also used an analytic expression to calculate the starting frequency of the waveform in the detector. The TS group expanded both PN amplitude series as was necessary, but did not truncate the ensuing ratio. They also used a numerical inversion of the PN series for $t(f)$ at $t = 0$ to find the starting frequency. Their code allowed for non-zero spins, but not for spin precession. For code comparisons, the spins were set to zero.

Table 3. Fractional differences in the claimed parameter-estimation accuracy between the AISSV and TS models. The parameters we consider are time to coalescence, chirp mass, luminosity distance, sky and orbital angular resolution, and orbital phase at coalescence. The angular resolution variables are defined by $\Delta\Omega_i = 2\pi[\Sigma^{\theta\theta}\Sigma^{\phi\phi} - (\Sigma^{\theta\phi})^2]^{1/2}$, where Σ^{ij} correspond to the elements of the variance-covariance matrix, and $i = S, L$. Note that again, values in the Table should be multiplied by 10^{-4} .

Case	Δt_c	$\Delta \ln(M_c)$	$\Delta \ln(D_L)$	$\Delta\Omega_S$	$\Delta\Omega_L$	$\Delta\varphi_c$
1a	3.87	16.40	8.35	1.04	0.46	13.18
2a	10.22	24.81	0.00	4.69	1.34	7.08

In Table 3 we compare the two groups' results. For this exercise, the AISSV group agreed to use the same PN truncation of the amplitude as the TS group. We can see that in all cases the results differ by less than 1%.

3.3. Comparison Between Time- and Frequency-Domain Waveforms

In initial comparisons between time- and frequency-domain codes we observed discrepancies in the SNRs. A visual comparison was made by inverse Fourier transforming the frequency domain waveforms and comparing with the time-domain waveforms. It turned out that the definition of the phase at coalescence is slightly different in the time and frequency domains. In order to match waveforms, a slight correction of the phase at coalescence was needed in the frequency-domain waveforms.

Another aspect affecting the results was the fact that both time-domain models used a truncation function to smoothly reduce the signal to zero at the end of the inspiral phase. As this truncation function was also affecting the comparison between the time- and frequency-domain approaches, it was decided that an upper frequency cutoff of 0.1 mHz would be applied to case 1a, and an upper frequency cutoff of 1 mHz would be applied to case 2a. This nullified the effect of the time-domain taper. In Table 4 we compare the four sets of results from the MM, AISSV, modified AISSV and TS entries. We can see that there is excellent agreement between all groups.

Table 4. Parameter-estimation accuracies for Case 2a for both time- and frequency-domain codes, assuming an upper frequency cutoff of 10^{-3} Hz. We present results for the MM, TS and AISSV groups. The AISSV* entry is the code modified to match the TS formalism.

Group	MM	TS	AISSV	AISSV*
SNR	101.2	100.9	99.41	100.8
$\Delta \ln(M_c)$	4.651×10^{-5}	4.671×10^{-5}	4.678×10^{-5}	4.659×10^{-5}
$\Delta t_c(\text{s})$	28.17	27.14	27.41	27.15
$\Delta \ln(D_L)$	0.0673	0.0674	0.0677	0.0675
$\Delta \Omega_S(\text{deg}^2)$	17.55	17.56	17.53	17.52
$\Delta \Omega_L(\text{deg}^2)$	17.05	17.13	17.57	17.29
$\Delta \varphi_c$	0.1225	0.1211	0.1213	0.1202

Finally, only the MM time-domain model had the ability to include at the same time higher harmonics, spins and spin precession. While no comparison was possible with the other groups taking part in this exercise, the MM code has been tested by comparison with an MLDC code written by Babak using the Kidder model to evolve spins [39, 43]. Because only the MM model had a working Fisher matrix, the comparison was done by generating waveforms for both models and then subtracting them from each other. This test yielded very small residuals and confirmed the validity of the MM treatment of spins and spin precession.

4. Results

In the foregoing sections we discussed the merger-tree models used in this work (Sec. 2) and described the cross-validation of the different parameter-estimation codes (Sec. 3). Here we will present the source distributions resulting from our merger-tree models, and use them for Monte Carlo simulations of LISA observations of MBHBs. One purpose of this exercise was to evaluate the impact of different design choices on LISA’s ability to accurately measure and localize MBHBs. Therefore, for each source distribution we computed parameter accuracies using two different LISA noise curves: the “baseline” and “6-link” noise models, described below.

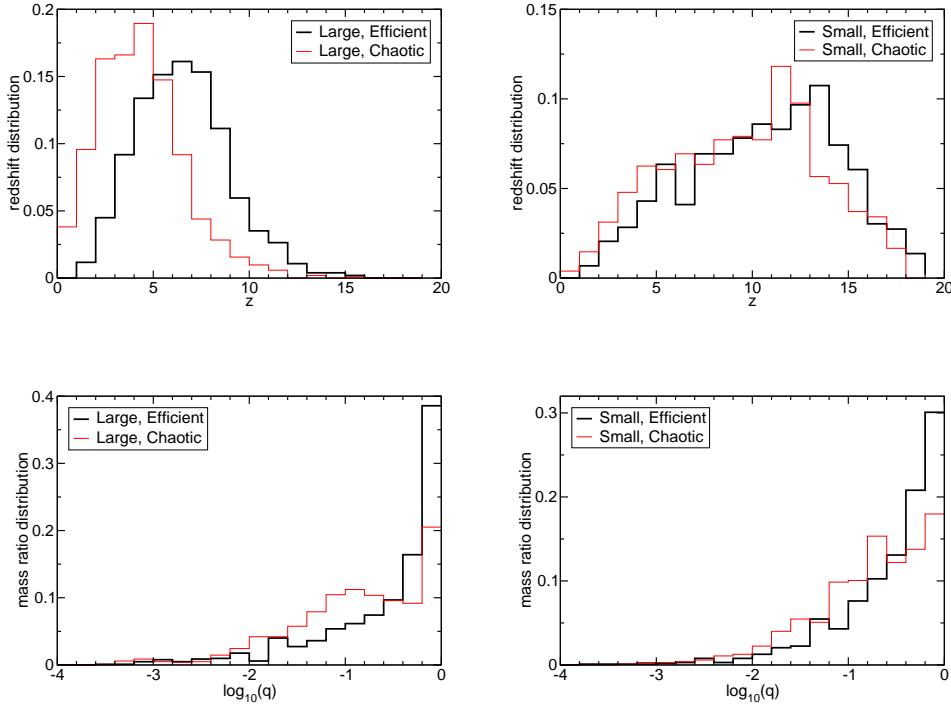


Figure 1. The distribution of MBHB mergers as a function of redshift z (top panels) and (logarithmic) mass ratio $\log(q)$ (bottom panels) for large-seed models (left) and small-seed models (right). Thick (black) histograms refer to efficient accretion, thin (red) histograms to chaotic accretion.

4.1. Source distributions

To assess what science LISA can do with MBHB mergers it is essential to understand the distribution of sources as a function of redshift, masses and spins of the component black holes. We evaluated parameter accuracies for the four different merger models described in Sec. 2. We shall call the models SE (small seeds, efficient accretion), SC (small seeds, chaotic accretion), LE (large seeds, efficient accretion) and LC (large seeds, chaotic accretion).

For large-seed models most mergers occur in the redshift range $3 \lesssim z \lesssim 8$, with a peak around $z \sim 5$ (Fig. 1, top left panel). In the case of small seeds, mergers are roughly uniform in z over the range $4 \lesssim z \lesssim 15$, with a peak around $z \sim 12$ (Fig. 1, top right panel). Large-seed models are likely to produce more symmetric binaries (which produce larger SNR for fixed total mass). In contrast, small-seed models lead to more asymmetric systems (see Fig. 1, bottom panels). Thus, although small-seed MBHB mergers could occur more frequently, a smaller fraction would be observed by LISA due to their smaller total mass and less symmetric mass ratios.

Spin precession usually improves parameter-estimation accuracies [5, 8]. LE models are the only ones that produce binaries in which both black holes generally have large spins. In SE models the spin is large for the more massive black hole,

but it is often negligible for the smaller hole. In chaotic accretion models (LC and SC) spins are always negligible. For a more detailed discussion of spin evolution and observational implications, see [27].

4.2. Baseline and 6-link noise models

To evaluate the impact of different LISA design choices on the missions science performance, for each source distribution we computed parameter accuracies using two different LISA configurations and associated noise curves: a “6-link” configuration, which allows the construction of all three independent TDI channels, and a “baseline” configuration of 4 links, producing a single Michelson channel. The instrumental noise \tilde{S}_n is similar for the baseline and 6-link configurations, save for the location of the low frequency “wall”, which is at 10^{-4} Hz for the baseline and 3×10^{-5} Hz for the 6-link model. The instrumental noise in the Michelson A and E channels is given by

$$\tilde{S}_n(f) = \frac{1}{12L^2} \left\{ (2+k)\tilde{S}_p + \frac{1+k+k^2}{(2\pi f)^4} S_a \left[1 + \left(\frac{10^{-4}}{f} \right) \right] \right\}, \quad (5)$$

where $k = \cos(f/f_*)$. Here $\tilde{S}_p = 3.24 \times 10^{-22} \text{ m}^2 \text{ Hz}^{-1}$ is slightly different from the value adopted in Eq. (3), corresponding to the baseline currently adopted by the LIST, and the second term in square parentheses is the pink acceleration noise level. The confusion noise estimates $\tilde{S}_{\text{conf}}(f)$ used in the analysis of MBHB merger trees, unlike those for the comparison exercise described in Sec. 3, did include interacting binaries. The confusion noise for the baseline is slightly higher than for the 6-link configurations as the latter is a factor of ~ 2 more sensitive, and has a five- versus three-year nominal lifetime. The total (instrumental plus confusion) noise spectral densities for these models are plotted in Figure 2.

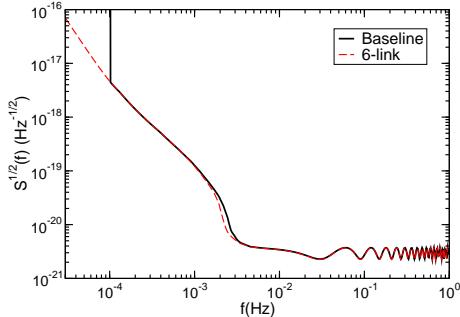


Figure 2. Baseline and 6-link noise models (including both instrumental and confusion noise) used in the merger-tree analysis.

4.3. Parameter accuracies

LISA’s parameter-estimation capabilities were calculated using the MM code to compute the covariance matrix. As discussed in Sec. 3, we have cross-validated their code against other codes finding excellent agreement. The advantage of the MM code

is that it incorporates not only high-order amplitude corrections, but also precessional effects due to spin. However, even the MM code does not take advantage of the merger and ringdown parts of the signal. Therefore, the parameter accuracies obtained in this work probably *underestimate* LISA's potential performance.

The statistics of detectability and parameter accuracies are based on a Monte Carlo simulation in which we randomly selected 1024 events with the distributions discussed in Sec. 4.1. We eliminated from the sample all binaries with total (locally-measured) mass larger than $3 \times 10^7 M_\odot$, since a) the code often crashed for such high masses, b) these cases represented a very small fraction of the population, and c) their inspirals take place almost entirely outside the LISA band, and so would be very poorly localized in any case. We then calculated 2048 covariance matrices (1024 for each of the two LISA noise models) on JPL's Cosmos supercomputer. About two percent of the jobs did not finish within a preset 3 hour period, and some 0.1% of the simulations had failures; such cases were simply discarded‡.

Table 5. For each merger-tree model (SE, SC, LE and LC) we list: the total number N of MBHB events in LISA's past light cone in a one-year observation; the number of events N_{det} detectable with SNR larger than 10 in one year; the number for which the error in the luminosity distance D_L is 10% or less; the number that is localizable within 1 and 10 deg 2 ($N_{1 \text{ deg}^2}$ and $N_{10 \text{ deg}^2}$, respectively); the number that can be resolved to within 10 deg 2 with less than 10% errors in D_L ($N_{10 \text{ deg}^2, 10\% D_L}$), and the same for 1 deg 2 in angular resolution and 1 % error in D_L ($N_{1 \text{ deg}^2, 1\% D_L}$). Results for the 6-link model are followed by those for the baseline model (within parentheses).

Model	N	N_{det}	$N_{10\% D_L}$	$N_{10 \text{ deg}^2}$	$N_{10 \text{ deg}^2, 10\% D_L}$	$N_{1 \text{ deg}^2}$	$N_{1 \text{ deg}^2, 1\% D_L}$
SE	80	33 (25)	21 (8.0)	8.2 (1.5)	7.9 (1.1)	2.2 (0.6)	1.7 (0.1)
SC	75	34 (27)	17 (4.4)	6.1 (0.4)	5.5 (0.4)	1.3 (0.1)	1.3 (0.1)
LE	24	23 (22)	21 (7.7)	10 (0.8)	10 (0.7)	2.2 (0.1)	1.2 (0.05)
LC	22	21 (19)	14 (4.3)	6.5 (0.5)	5.4 (0.5)	1.8 (0.04)	1.0 (0.1)

For our four merger-tree models (as listed in column 1 of Table 5) we first give the total number N of mergers in LISA's past light-cone during a one-year observation (column 2) and the number of events N_{det} that are detectable in one year with SNR larger than 10 (column 3). The rest of the columns show the number of observable events with: error in D_L of 10% or less (column 4); angular resolution $\Delta\Omega_S < 10 \text{ deg}^2$ (column 5); $\Delta\Omega_S < 1 \text{ deg}^2$ and error in D_L of 10% or less (column 6); $\Delta\Omega_S < 1 \text{ deg}^2$ (column 7); $\Delta\Omega_S < 1 \text{ deg}^2$ and error in D_L of 1% or less (column 8).

LISA will detect quite a good fraction of the mergers which occur in the universe: almost all mergers will be detected in large-seed scenarios, and nearly half of all mergers in small-seed scenarios. This is because large seed black holes lead to more massive MBHBs with mass ratio q close to one, so they can be seen out to larger redshifts. Since the number of mergers in LISA's past light cone is three times larger in the case of small seeds, the number of detectable events will be similar in both cases. Table 5 shows that the number of detectable events (about 30 per year in small-seed

‡ Our covariance matrices are computed in 15 dimensions. Because of large covariances between parameters (often close to ± 1) the Fisher information matrix is ill-conditioned, and we believe our occasional failures to arise when numerical inversion of the Fisher matrix fails within the available numerical precision. We will explore the cause in more detail in the future.

models, and about 20 per year in large-seed models) seems to be quite insensitive to the details of accretion and spin evolution. The number of detectable events is slightly smaller for the baseline model than for the 6-link model.

Measurements of D_L accurate within 1% would allow us to determine the dark energy equation-of-state parameter w to better than 1%, but this would not be possible if the accuracy in D_L were of order 10% [10]. The typical error in D_L due to weak lensing is in the range 5–10% at $z \sim 2$ [44, 45]. Ideal MBHB sources to do precision cosmology (“standard sirens”) would be in the redshift range (say) $0.5 < z < 0.8$, where lensing errors would presumably be of order 2–3%. Therefore, an accuracy in D_L of about 1% is needed if we want our standard sirens to be limited only by lensing, and not by the random errors due to LISA’s instrumental noise. The temporal coincidence with any prompt electromagnetic counterpart may offer the best chance of identifying a unique host galaxy, especially if we can locate the source within one (or a few) square degrees [46].

Table 5 clearly shows that with the merger-tree models and gravitational waveforms used in this study, both the baseline design and the 6-link model have the potential for precision cosmology. However, the 6-link configuration performs significantly better than the baseline design. With the 6-link configuration LISA can get the luminosity distance to within 10% for the majority of detectable events. This is not really surprising, since the luminosity distance does not have a strong correlation with the phasing parameters and our detection threshold is $\rho_{\text{th}} = 10$. However, only a fourth to half of all events can be resolved within a 10 deg^2 error box. Once an event has a good sky resolution it will also have a good accuracy in D_L , as evidenced by the similarity of the numbers in columns 5 and 6: this is due to the high correlation between luminosity distance and angular parameters of the source (see e.g. [6, 8]).

The fraction of events that can be localized to within 1 deg^2 is far smaller. A few events can be localized to such accuracy using one-year observations and the 6-link noise curve, while the statistics are too low to determine if events could be localized to such accuracy with the baseline design. Finally, while we have considered only 6-link and 4-link LISA configurations, it is possible that LISA will operate with 5 links for some substantial time. The important point is that for MBHB mergers, LISAs sensitivity and parameter extraction accuracy with 5 links should be nearly as good as with 6 [47]. Therefore it is quite plausible that ~ 1 event per year could have angular resolution and accuracy in luminosity distance good enough to measure the dark-energy equation-of-state parameter w [10].

5. Conclusions

We compared four largely independent codes for calculating LISA’s parameter-estimation capabilities. All codes are based on the Fisher-matrix approximation, but in the past they used somewhat different signal models, source parametrizations and noise curves. We showed that once these differences are removed, the four codes give results in extremely close agreement with each other.

Using a code that includes both spin precession and higher harmonics in the GW signal, we presented a preliminary exploration of LISA’s ability to do precision cosmology using four different merger-tree models and two different sensitivity curves. For the merger history of galaxies we used either “small” seed black holes ($m_{\text{seed}} \sim 100 M_\odot$) or “large” seeds ($m_{\text{seed}} \sim 10^5 M_\odot$). In each case we adopted two different accretion scenarios (efficient or chaotic accretion), thereby giving us four models for

the birth, growth and spin evolution of massive black holes. We then employed a carefully tested tool to compute the parameter accuracies for events that LISA can detect with an (amplitude) SNR of at least 10.

Our study shows that LISA has good potential for carrying out precision cosmology. There will be about 20 sources with a modest distance measurement (to within 10 %) and about 10 sources with a modest sky resolution (10 deg^2). Even more interestingly, each year LISA may observe a few sources with excellent sky resolution (1 deg^2) and luminosity distance measurements (to within 1%). Such accuracies are good enough to measure the dark-energy equation-of-state parameter w to within a percent, the dominant source of uncertainty being errors in the luminosity distance from weak lensing of GW events [44, 45].

The waveforms used in our study included important physical effects, such as high-order PN corrections in amplitude and phase and precessing spins. However our waveforms were incomplete in one respect: they did not include the merger and ringdown parts of the signal. Future work will include ringdown either in a phenomenological way (by estimating the final black hole's parameters from the inspiral parameters, and then using the ringdown measurement formalism developed in [48]) or by “stitching” PN inspiral waveforms to a set of numerical relativity waveforms. As shown recently by Babak *et al.* [49] (see also [50]) the inclusion of the merger and ringdown signal is expected to improve parameter estimation in two ways: by enhancing the signal-to-noise ratio and by increasing the signal bandwidth. Future studies will need to explore the contribution of merger and ringdown signals to detection rates and parameter-estimation accuracy. They should also explore a broader class of merger-tree models to get a more realistic evaluation of LISA's ability to do precision cosmology.

The LISAPE Taskforce is presently developing tools similar to those described in this paper to explore measurement accuracies from extreme mass-ratio inspirals (EMRIs) of compact objects into massive black holes. This work will be reported elsewhere.

6. Acknowledgments

We wish to thank Michele Vallisneri for very helpful interactions. E.B.'s and C.C.'s work was carried out at the Jet Propulsion Laboratory (JPL), California Institute of Technology, under contract with the National Aeronautics and Space Administration. C.C.'s work was partly supported by NASA Grant NNX07AM80G. E.B.'s research was supported by an appointment to the NASA Postdoctoral Program at JPL, administered by Oak Ridge Associated Universities through a contract with NASA. N.C. is supported by NASA grant NNX07AJ61G. A.S. and M.T. would like to thank the support of the Max-Planck Society, the Spanish Ministerio de Educación y Ciencia Research Projects FPA-2007-60220, HA2007-0042, CSD207-00042 and the Govern de les Illes Balears, Conselleria d'Economia, Hisenda i Innovació. S.A.H and R.N.L. have been supported by NASA Grants NNG05G105G and NNX08AL42G, as well as NASA Contract No. 1291617 and the MIT Class of 1956 Career Development Fund. I.M. was partially supported by NASA ATP Grant NNX07AH22G to Northwestern University. E.K.P. would like to thank the DLR (Deutsches Zentrum für Luft- und Raumfahrt) for support during this work. Research at Cardiff was supported in part by PPARC grant PP/F001096/1. The Monte Carlo parameter-estimation results presented here were generated on JPL's Cosmos supercomputer. ©2008. All rights reserved.

- [1] C. Cutler, Phys. Rev. D **57**, 7089 (1998).
- [2] T. A. Moore and R. W. Hellings, Phys. Rev. D **65**, 062001 (2002).
- [3] A. M. Sintes and A. Vecchio, arXiv:gr-qc/0005058.
- [4] R. W. Hellings and T. A. Moore, Class. Quant. Grav. **20**, S181 (2003).
- [5] A. Vecchio, Phys. Rev. D **70**, 042001 (2004).
- [6] E. Berti, A. Buonanno and C. M. Will, Phys. Rev. D **71**, 084025 (2005).
- [7] E. Berti, A. Buonanno and C. M. Will, Class. Quant. Grav. **22**, S943 (2005).
- [8] R. N. Lang and S. A. Hughes, Phys. Rev. D **74**, 122001 (2006) [Erratum-ibid. D **75**, 089902 (2007); Erratum-ibid. D **77**, 109901 (2008)].
- [9] K. G. Arun, B. R. Iyer, B. S. Sathyaprakash and S. Sinha, Phys. Rev. D **75**, 124002 (2007).
- [10] K. G. Arun, B. R. Iyer, B. S. Sathyaprakash, S. Sinha and C. V. D. Broeck, Phys. Rev. D **76**, 104016 (2007) [Erratum-ibid. D **76**, 129903 (2007)].
- [11] M. Trias and A. M. Sintes, Phys. Rev. D **77**, 024030 (2008).
- [12] R. N. Lang and S. A. Hughes, Astrophys. J. **677**, 1184 (2008).
- [13] E. K. Porter and N. J. Cornish, Phys. Rev. D **78**, 064005 (2008).
- [14] M. Trias and A. M. Sintes, Class. Quant. Grav. **25**, 184032 (2008).
- [15] The LISA Science Team, http://www.srl.caltech.edu/lisa/mission_documents.html
- [16] E. Berti, Class. Quant. Grav. **23**, S785 (2006) [arXiv:astro-ph/0602470].
- [17] A. Sesana, M. Volonteri and F. Haardt, Mon. Not. Roy. Astron. Soc. **377**, 1711 (2007).
- [18] M. Volonteri, arXiv:0709.1722 [astro-ph].
- [19] S. Babak *et al.*, Class. Quant. Grav. **25**, 184026 (2008).
- [20] M. Vallisneri, Phys. Rev. D **77**, 042001 (2008).
- [21] M. Volonteri, F. Haardt and P. Madau, Astrophys. J. **582**, 559 (2003).
- [22] M. C. Begelman, M. Volonteri and M. J. Rees, Mon. Not. Roy. Astron. Soc. **370**, 289 (2006).
- [23] J. M. Bardeen, Nature **226**, 64 (1970).
- [24] K. S. Thorne, Astrophys. J. **191**, 507 (1974).
- [25] A. R. King and J. E. Pringle, Mon. Not. Roy. Astron. Soc. Lett. **373**, L93 (2006).
- [26] S. A. Hughes and R. D. Blandford, Astrophys. J. **585**, L101 (2003) [arXiv:astro-ph/0208484].
- [27] E. Berti and M. Volonteri, Astrophys. J. **684**, 822 (2008).
- [28] L. Rezzolla, E. Barausse, E. N. Dorband, D. Pollney, C. Reisswig, J. Seiler and S. Husa, Phys. Rev. D **78**, 044002 (2008).
- [29] T. Bogdanovic, C. S. Reynolds and M. C. Miller, Astrophys. J. Lett. **661**, 147 (2007).
- [30] C. G. Lacey and S. Cole, Mon. Not. Roy. Astron. Soc. **262**, 627 (1993).
- [31] W. H. Press and P. Schechter, Astrophys. J. **187**, 425 (1974).
- [32] R. K. Sheth and G. Tormen, Mon. Not. Roy. Astron. Soc. **329**, 61 (2002).
- [33] M. G. Haehnelt, Mon. Not. Roy. Astron. Soc. **269**, 199 (1994).
- [34] S. Babak *et al.* [Mock LISA Data Challenge Task Force Collaboration], Class. Quant. Grav. **25**, 114037 (2008).
- [35] G. Nelemans, L. R. Yungelson and S. F. Portegies Zwart, Astron. Astrophys. **375**, 890 (2001).
- [36] G. Nelemans, L. R. Yungelson and S. F. Portegies Zwart, Mon. Not. Roy. Astron. Soc. **349**, 181 (2004).
- [37] N. J. Cornish and T. B. Littenberg, Phys. Rev. D **76**, 083006 (2007).
- [38] L. Blanchet, B. R. Iyer, C. M. Will and A. G. Wiseman, Class. Quant. Grav. **13**, 575 (1996).
- [39] LISATools website and SVN repository, <http://lisatools.googlecode.com>
- [40] T. A. Apostolatos, C. Cutler, G. J. Sussman and K. S. Thorne, Phys. Rev. D **49**, 6274 (1994).
- [41] K. A. Arnaud *et al.*, AIP Conf. Proc. **873**, 619 (2006).
- [42] K. S. Thorne, in 300 Years of Gravitation, eds. S. W. Hawking and W. Israel (Cambridge University Press, Cambridge, UK, 1987), p. 330.
- [43] L. E. Kidder, Phys. Rev. D **52**, 821 (1995).
- [44] D. E. Holz and S. A. Hughes, Astrophys. J. **629**, 15 (2005) [arXiv:astro-ph/0504616].
- [45] B. Kocsis, Z. Frei, Z. Haiman and K. Menou, Astrophys. J. **637**, 27 (2006) [arXiv:astro-ph/0505394].
- [46] B. Kocsis, Z. Haiman and K. Menou, Astrophys.J. 684 (2008) 870-888
- [47] M. Vallisneri, J. Crowder and M. Tinto, Class. Quant. Grav. **25**, 065005 (2008) [arXiv:0710.4369 [gr-qc]].
- [48] E. Berti, V. Cardoso and C. M. Will, Phys. Rev. D **73**, 064030 (2006).
- [49] S. Babak, M. Hannam, S. Husa and B. F. Schutz, arXiv:0806.1591 [gr-qc].
- [50] J. I. Thorpe, S. T. McWilliams, B. J. Kelly, R. P. Fahey, K. Arnaud and J. G. Baker, arXiv:0811.0833 [astro-ph].