

Report on the second Mock LISA Data Challenge

The Mock LISA Data Challenge Task Force: Stanislav Babak¹, John G. Baker², Matthew J. Benacquista³, Neil J. Cornish⁴, Jeff Crowder⁵, Curt Cutler^{5,6}, Shane L. Larson⁷, Tyson B. Littenberg⁴, Edward K. Porter¹, Michele Vallisneri^{5,6}, Alberto Vecchio^{8,9} and *the Challenge-2 participants:* Gerard Auger¹⁰, Leor Barack¹¹, Arkadiusz Błaut¹², Ed Bloomer¹³, Duncan A. Brown^{14,15,6}, Nelson Christensen¹⁶, James Clark¹³, Stephen Fairhurst^{17,15,6}, Jonathan R. Gair¹⁸, Hubert Halloin¹⁰, Martin Hendry¹³, Arturo Jimenez³, Andrzej Królak¹⁹, Ilya Mandel^{9,6}, Chris Messenger¹³, Renate Meyer²⁰, Soumya Mohanty³, Rajesh Nayak³, Antoine Petiteau¹⁰, Matt Pitkin¹³, Eric Plagnol¹⁰, Reinhard Prix¹, Emma L. Robinson⁸, Christian Roever²⁰, Pavlin Savov⁶, Alexander Stroer^{8,9}, Jennifer Toher¹³, John Veitch⁸, Jean-Yves Vinet²¹, Linqing Wen¹, John T. Whelan¹, Graham Woan¹³

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

² Gravitational Astrophysics Lab., NASA Goddard Space Flight Center, 8800 Greenbelt Rd., Greenbelt, MD 20771, USA

³ Center for Gravitational Wave Astronomy, Univ. of Texas at Brownsville, Brownsville, TX 78520, USA

⁴ Dept. of Physics, Montana State Univ., Bozeman, MT 59717, USA

⁵ Jet Propulsion Laboratory, California Inst. of Technology, Pasadena, CA 91109, USA

⁶ Theoretical Astrophysics, California Inst. of Technology, Pasadena, CA 91125

⁷ Dept. of Physics, Weber State Univ., 2508 University Circle, Ogden, UT 84408, USA

⁸ School of Physics and Astronomy, Univ. of Birmingham, Edgbaston, Birmingham B152TT, UK

⁹ Dept. of Physics and Astronomy, Northwestern Univ., Evanston, IL, USA

¹⁰ APC, UMR 7164, Univ. Paris 7 Denis Diderot, 10, rue Alice Domon et Leonie Duquet, 75025 Paris Cedex 13, France

¹¹ School of Mathematics, Univ. of Southampton, Southampton, SO171BJ, UK

¹² Inst. of Theoretical Physics, Univ. of Wrocław, Wrocław, Poland

¹³ Dept. of Physics and Astronomy, Univ. of Glasgow, Glasgow, UK

¹⁴ Dept. of Physics, Syracuse Univ., Syracuse, NY13244, USA

¹⁵ LIGO Laboratory, California Inst. of Technology, Pasadena, CA 91125

¹⁶ Physics and Astronomy, Carleton College, Northfield, MN, USA

¹⁷ School of Physics and Astronomy, Cardiff Univ., 5, TheParade, Cardiff, UK, CF243YB

¹⁸ Inst. of Astronomy, Univ. of Cambridge, Cambridge, CB30HA, UK

¹⁹ Inst. of Mathematics, Polish Academy of Sciences, Warsaw, Poland

²⁰ Dept. of Statistics, The Univ. of Auckland, Auckland, New Zealand

²¹ ARTEMIS, Observatoire de la Cote d'Azur-C.N.R.S., 06304 Nice, France

Abstract. The Mock LISA Data Challenges are a program to demonstrate LISA data-analysis capabilities and to encourage their development. Each round of challenges consists of several data sets containing simulated instrument noise and gravitational-wave sources of undisclosed parameters. Participants are asked to analyze the data sets and report the maximum information about source parameters. The challenges are being released in rounds of increasing complexity and realism: in this proceeding we present the results of Challenge 2, issued in January 2007, which successfully demonstrated the recovery of signals from supermassive black-hole binaries, from 20,000 overlapping Galactic white-dwarf binaries, and from the extreme-mass-ratio inspirals of compact objects into central galactic black holes.

PACS numbers: 04.80.Nn, 95.55.Ym

1. Introduction

The Laser Interferometer Space Antenna (LISA), a NASA and ESA space mission to detect gravitational waves (GWs) in the 10^{-5} – 10^{-1} Hz range [1], will produce time series consisting of the superposition of the signals from millions of sources, from our Galaxy to the edge of the observable universe. Some of the signals (such as those from extreme-mass-ratio inspirals, or EMRIs) are very complex functions of the physical parameters of the sources; others (such as those from Galactic white-dwarf binaries) are simpler, but their resolution will be confused by the presence of many other similar signals overlapping in frequency space. Thus, data analysis is integral to the LISA measurement concept, because no source can be identified without first carefully teasing out its individual voice from the noisy party of each data set. Understanding data analysis is therefore important to demonstrate that LISA can meet its science requirements, and to translate them into decisions about instrument design.

The idea of the Mock LISA Data Challenges (MLDCs) arose in late 2005 from this realization. The MLDCs have the purpose of encouraging and tracking progress in LISA data-analysis development, and (as a useful byproduct) of prototyping the LISA computational infrastructure: common data formats, standard models of the LISA orbits, noises and measurements, software to generate waveforms and to simulate the LISA response, and more. The MLDCs are a coordinated (but voluntary) effort in the GW community, whereby a task force chartered by the LISA International Science Team periodically issues a number of data sets containing synthetic noise and GW signals from sources of undisclosed parameters; challenge participants return detection candidates and parameter estimates, together with descriptions of their search methods. These results are then compiled and compared to the previously secret challenge “key”.

Challenge 1, issued in Jun 2006 with results due in Dec 2006 (see [2, 3]), tackled the detection and parameter characterization of *verification binaries* (Galactic binaries of known frequency and position); of loud unknown Galactic binaries, either alone or in small, moderately interfering groups; and of relatively loud inspirals of supermassive-black-hole (MBH) binaries. All sources were represented by somewhat idealized waveforms, and they were staged on instrument noise alone. Altogether, Challenge 1 successfully demonstrated the detection of all three source classes. Ten research collaborations submitted entries, adopting a variety of methods (template-bank, stochastic and genetic matched filtering; time-frequency; tomography; Hilbert transform). Despite the short timescale, all challenges were “solved” by at least one

group, although some searches locked on strong secondary maxima of the source parameter probabilities. More important, Challenge 1 helped set the playing field and assemble the computational tools for the more realistic Challenge 2.

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

Thirteen collaborations (comprising all the researchers listed as participants in the byline of this article, and most task force members) submitted a total of 22 entries, including a proof-of-principle analysis for stochastic backgrounds performed on data set 2.1. Altogether, Challenge 2 successfully demonstrated the identification of $\sim 20,000$ Galactic binaries, the accurate estimation of MBH inspiral parameters, and the positive detection of EMRIs. In the rest of this paper, we describe some of its highlights. All the solutions submitted by participating groups, together with technical write-ups of their methods and findings, can be found at the URL www.tapir.caltech.edu/~mldc/results2. A few groups are also contributing descriptions of their work to the proceedings of this conference.

2. Data sets 2.1 and 2.2: The Galaxy

Five groups submitted Galactic-binary catalogs for data sets 2.1 and 2.2:

GLIG A collaboration of research groups at institutions in the UK, United States and New Zealand developed a Reversible-Jump Markov Chain Monte Carlo (RJ MCMC) code that can sample models with different numbers of sources; for lack of time, however, they only submitted parameter sets for the verification binaries.

IMPAN Królak and Blaut developed an \mathcal{F} -statistic, template-bank-based matched-filtering search [5, 6], and submitted parameters for 404 sources for data set 2.1.

MTJPL The Montana State–JPL collaboration used a Metropolis–Hastings Monte Carlo (MHMC) code that ran separately for overlapping frequency bands and for different hypothesized numbers of sources; model comparison was then used to determine the most probable number of sources in each band. The collaboration submitted parameter sets for 19,324 sources for data set 2.1, and 18,461 sources for data set 2.2.

PrixWhelanAEI Prix and Whelan developed an \mathcal{F} -statistic, template-bank-based matched-filtering search using a hierarchical scheme that enforced trigger coincidence between TDI observables, followed by a coherent follow-up using noise-orthogonal TDI combinations [11]. They submitted parameter sets for 1777 sources for data set 2.1, and 1737 sources for data set 2.2.

UTB Nayak, Jimenez and Mohanty used a tomographic reconstruction technique and submitted parameters for 3862 sources in data set 2.1.

Evaluating the performance of these searches brings up several problems of principle: while we know that many of the ~ 30 million Galactic sources that were injected

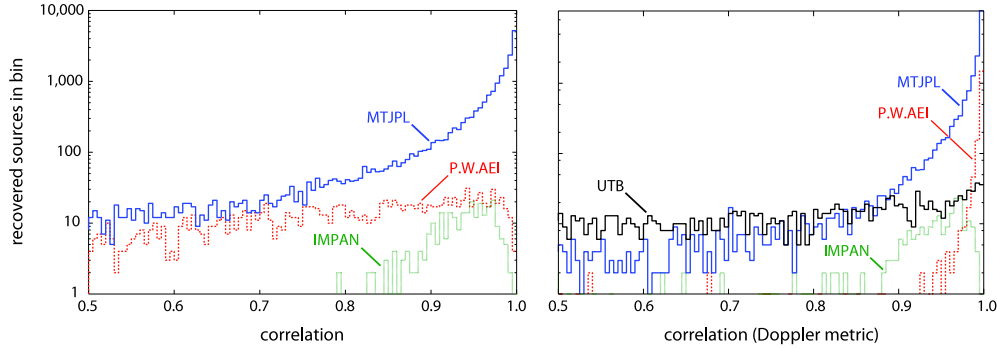


Figure 1. Correlation analysis of Challenge-2.1 Galactic-binary catalogs (histogram, with bin fractions given on logarithmic scale). Left panel: associated by correlation; right panel: associated by Doppler metric.

into the data sets cannot be recovered because they are (relatively) too weak, we do not have a precise estimate of how many sources should be recoverable. Thus, the notion of false dismissal is not well defined. To make matters worse, the notion of false positive is also ill defined, because a single recovered source can provide a good fit to the blended signal from several injected sources, which may well be the “right” answer with the knowledge we have, since it is the best fit to the data with the smallest numbers of parameters.

The task force devoted considerable time to the analysis of Galactic-binary searches, and we do not have space here to describe all the treatments that we applied to the data. Instead we will limit our report to techniques that pair up individual recovered sources with individual sources from the challenge key, with the understanding that this will overestimate the number of false positives, and say nothing about false dismissals.

One way to proceed is to associate the reported and injected sources that have the strongest signal correlation, limiting the search to the bright ($\text{SNR} > 2$) injected sources that could in principle have been found: in the left panel of figure 1 we show the distribution of correlations generated with this procedure for data set 2.1. Detections with the highest correlations can be considered “safest,” while those with the lowest correlations probably represent spurious associations.

Another procedure is to associate the reported and injected sources that minimize the *Doppler metric* that spans the frequency–sky–location subspace of the full parameter space, and automatically maximizes correlation over the *extrinsic parameters* (amplitude, polarization, inclination, initial phase): the right panel of figure 1 shows the resulting distribution of correlations. The UTB entry, which includes frequency and sky position but not the extrinsic parameters, can only be plotted this way. Generally, this is a softer criterion, and all searches do better by it (especially the PrixWhelanAEI entry, whose long-wavelength approximation for the LISA response is prone to extrinsic-parameter errors).

Figure 2 shows the SNRs of the recovered sources and the errors for the intrinsic and extrinsic parameters, computed after associating sources by correlation (and by Doppler metric only for the UTB entry), again for data set 2.1. The error in frequency is in most cases within a small fraction of a Fourier bin, and the errors in sky position are within a few degrees; by contrast, the errors in the amplitude and

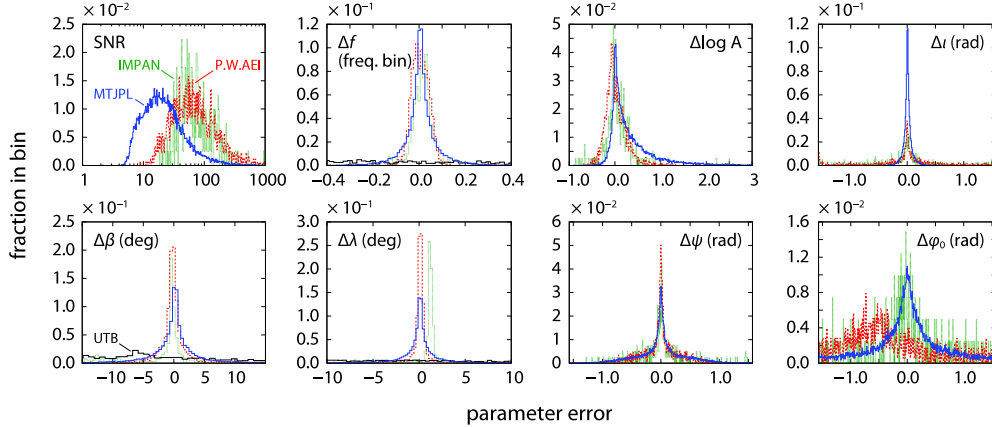


Figure 2. Recovered SNRs and intrinsic and extrinsic parameter errors for Challenge-2.1 Galactic-binary catalogs (histogram). True sources and templates are associated by correlation, except for the UTB catalog, for which they are associated by Doppler metric.

in the (extrinsic) orientation angles are larger. The ϕ_0 graph for the PrixWhelanAEI suggests a systematic error in the definition of initial phase.

Altogether, these challenges demonstrated a solid capability in analyzing signals from the Galaxy and resolving a large number of binaries. As we mentioned, deciding how well they were recovered is not an easy question to answer, because of the difficulty of defining (at least operationally) a notion of *identity* for recovered sources. These problems deserve careful attention in the future.

3. Data set 2.2: MBH binaries (over the Galaxy)

Four groups reported parameter sets for the MBH binaries in data set 2.2:

AEIse Babak and Porter used an \mathcal{F} -statistic, template-bank-based matched-filtering search, followed by an MCMC stage.

MTAEI Cornish and Porter used an MHMC matched-filtering search with a *frequency-annealed* scheme where shorter, lower-frequency templates are used in the initial phases of the search and then progressively extended.

JPLCT The JPL-CIT collaboration used a three-stage pipeline consisting of a track search in the time-frequency (TF) plane, followed by template-bank-based matched filtering, and by an MCMC refinement [8].

LisaFrance The French collaboration used a TF track search alone, and therefore reported only mass and time-of-coalescence parameters.

All the four MBH binaries in the data set (MBH-1, 2, 4 and 5, with total SNRs $\sim 2583, 25, 174$ and 117) were positively detected by AEIse, MTAEI, and JPLCT; the TF method used by LisaFrance identified MBH-1 and 4, but not MBH-2, and could report only a time of coalescence for MBH-5.

Tables 1 and 2 show fractional parameter errors for MBH-1 and 4, together with the SNR recovered by the best-fit candidates, computed as

$$\text{SNR}_{\text{best}} = \frac{(A_{\text{true}}|A_{\text{best}}) + (E_{\text{true}}|E_{\text{best}})}{\sqrt{(A_{\text{best}}|A_{\text{best}}) + (E_{\text{best}}|E_{\text{best}})}}, \quad (1)$$

Table 1. Recovered SNR and parameter errors for MBH-1. All parameters are defined as in tables 2 and 4 of [4], except for $\mu = m_1 m_2 / (m_1 + m_2)$, $M_c = (m_1 m_2)^{3/5} / (m_1 + m_2)^{1/5}$, and ϕ_c defined as the GW phase at coalescence. All errors on angles are given in radians; the true (optimal) SNR is **2583.42**.

	SNR	$\Delta M_c / M_c$	$\Delta \mu / \mu$	$\Delta t_c / t_c$	$\Delta \beta$	$\Delta \lambda$	$\Delta D / D$	$\Delta \iota$	$\Delta \psi$	$\Delta \phi_c$
		$\times 10^{-5}$	$\times 10^{-5}$	$\times 10^{-6}$	$\times 10^{-1}$	$\times 10^{-2}$	$\times 10^{-2}$	$\times 10^{-2}$	$\times 10^{-1}$	$\times 10^{-1}$
AEIse	2247.60	147.0	3386.2	19.0	-5.07	-82.1	77.8	-6.86	13.8	-6.82
MTAEI	2583.34	7.9	9.1	3.6	1.65	-1.2	3.5	4.94	1.2	-7.70
JPLCT	2582.42	27.5	28.7	16.0	4.81	12.2	12.3	-3.19	-12.1	7.45
lisaFrance	-	2944.6	72.7	67.8	-	-	-	-	-	-

Table 2. Recovered SNR and parameter errors for MBH-4, given as in table 1. The true (optimal) SNR is **174.12**.

	SNR	$\Delta M_c / M_c$	$\Delta \mu / \mu$	$\Delta t_c / t_c$	$\Delta \beta$	$\Delta \lambda$	$\Delta D / D$	$\Delta \iota$	$\Delta \psi$	$\Delta \phi_c$
		$\times 10^{-6}$	$\times 10^{-4}$	$\times 10^{-6}$	$\times 10^{-2}$	$\times 10^{-2}$	$\times 10^{-3}$	$\times 10^{-3}$	$\times 10^{-3}$	$\times 10^{-1}$
AEIse	81.38	1396.3	149.9	3.4	-12.5	104.7	574.0	3.5	-185.4	8.1
MTAEI	174.13	148.8	21.3	2.1	2.4	2.1	15.1	2.8	-7.5	1.7
	174.11	17.1	20.5	33.3	-42.4	-310.6	16.7	-13.7	-146.3	-6.3
JPLCT	174.11	4.2	9.5	2.1	7.8	9.6	1.2	1.3	-21.3	-5.4
	174.12	124.7	9.0	35.4	-47.3	-302.9	6.3	-12.4	1436.4	-12.4
lisaFrance	-	34394.1	1804.1	280.8	-	-	-	-	-	-

with $(\cdot|\cdot)$ the usual noise-weighted inner product, and $A = (2X - Y - Z)/3$ and $E = (Z - Y)/\sqrt{3}$ two noise-orthogonal TDI observables (see, e.g., [9]). In table 1, we see that the JPLCT search for MBH-1 locked onto a secondary probability maximum with SNR_{best} only slightly lower than the optimal value, but with sky positions off by several degrees, which also led to errors in the other parameters. The JPLCT authors hypothesize that this was caused by first subtracting a rough MBH-1 model from the data, then subtracting the resolvable Galactic binaries, and finally refining the MBH search.

MBH-4 (table 1) is an interesting example of a “true” bimodal probability distribution for the source parameters. MTAEI and JPLCT each submitted two candidates, placed at rather different sky locations, quoting relative probability ratios of 1:1 and 1.18:1. In this case, it was probably the sky position and orientation of this source that conspired to degrade LISA’s positional sensitivity, since they resulted in a very weak signal in one of the noise-orthogonal observables.

Altogether, this challenge demonstrated a solid capability in the detection and parameter estimation of MBH inspirals with moderate SNRs, even in the presence of a strong Galactic background, at least if the inspirals can be considered close to our idealized model: circular and adiabatic with negligible spin effects. These restrictions are being relaxed for the upcoming Challenge 3.

Table 3. Recovered SNRs and parameter errors for the EMRI signal in data set 1.3.1. All errors are given as *fractions of the allowed prior range* for the corresponding parameters (0.15 for e_0), except for the errors on ν_0 and D . Not all parameters are shown. For their definitions, see tables 2 and 5 of [4]. The true (optimal) SNR is **130.98**.

	SNR	$\delta\beta$	$\delta\lambda$	$\delta\theta_K$	$\delta\phi_K$	δa	$\delta\mu$	δM	$\frac{\Delta\nu_0}{\nu_0}$	δe_0	$\frac{\Delta D}{D}$
BBGP	74.86	-0.33	-0.0095	-0.13	-0.076	0.28	-0.15	-0.51	0.017	0.21	-1.21
	72.96	-0.32	0.011	-0.15	-0.078	0.27	-0.15	-0.51	0.017	0.21	-1.22
	72.52	-0.28	0.025	-0.063	-0.036	0.41	-0.17	-0.35	-0.009	0.29	-2.15
	72.49	-0.28	0.025	-0.063	-0.034	0.41	-0.17	-0.36	-0.009	0.29	-2.17
	70.59	-0.31	-0.020	-0.36	-0.21	0.44	-0.12	-0.12	-0.03	0.28	-0.91
EtfAG	-	0.016	0.0012	-	-	-0.082	0.10	-0.17	0.0026	0.098	-
MT	74.85	0.15	0.47	-0.069	-0.15	-0.026	0.073	0.18	0.00025	-0.11	-0.71
	76.52	0.084	-0.49	-0.33	-0.10	-0.022	0.046	0.16	0.00026	-0.10	-0.70

4. Data sets 1.3.X: EMRIs

Three groups reported parameter sets for the EMRIs in data sets 1.3.1–1.3.4. No group tackled the problem of detecting these systems in data set 2.2 (on top of the Galactic background).

BBGP Babak and colleagues used an MCMC matched-filtering search that modeled the signal with a sequence of progressively longer templates (a *time-annealed* scheme).

EtfAG Gair, Mandel and Wen used a TF track search that (for now) targeted only the intrinsic parameters and sky position [10].

MT Cornish used an MHMC matched-filtering search, running it in parallel on individual month-long segments, which were subsequently strung together for full detections.

Table 3 shows typical recovered SNRs and errors. While it is clear that the matched-filtering searches locked on several secondary probability maxima with comparable probabilities, the recovered SNRs correspond to solid detections with exceedingly low false-alarm probabilities. The errors are quoted as fractions of the allowed parameter ranges, and they are quite large. Intriguingly, the TF search was the most accurate in determining the sky position. Altogether, these challenges demonstrated a positive capability of detecting EMRIs, at least if their signals are similar in complexity to the *kludge* waveforms used in this challenge [4]; however, the prospects for accurate parameter estimation are still uncertain, and a good focus for further challenges.

5. Conclusion

We are very excited about the outcome of the first two MLDCs, which have given a convincing demonstration that a significant portion of the LISA science objectives could already be achieved with techniques that are currently in hand. Most of the research groups that participated in Challenge 1 have successfully made the transition to the greater complexity of Challenge 2. Challenge 3 will continue to move in the direction of more realistic signals, featuring chirping Galactic binaries and precessing

binaries of spinning MBHs. It will also include two new classes of signals: an isotropic primordial GW background and bursts from the cusps of cosmic strings. Between July and Dec 2007 we are also running Challenge 1B, a repeat of Challenge 1 conceived to provide a softer entry point for research groups new to the MLDCs.

The MLDC conventions, file formats, and software tools (see lisatools.googlecode.com) have matured to the point where interested parties can use them to generate a variety of data sets. This enables a wealth of interesting side investigations, such as the studies of the LISA science reach that are now being undertaken by the LISA Science Team. To obtain more information and to participate in the MLDCs, see the official MLDC website (astrogravs.nasa.gov/docs/mldc) and the task force wiki (www.tapir.caltech.edu/listwg1b).

Acknowledgments

SB, EKP, and JTW's work was supported by the German Aerospace Center (DLR) and the Max-Planck Society. MB acknowledges funding from NASA Grant NNG04GD52G and support from the NASA Center for Gravitational Wave Astronomy at University of Texas at Brownsville (NAG5-13396). NC was supported by NASA Grants NNG05GI69G and NNX07AJ61G. DB and SF acknowledge funding from NSF grant PHY-0601459 and the LIGO Laboratory. JC's, CC's and MV's work was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. JG acknowledges support from St Catharine's College, Cambridge. IM would like to thank the Brinson Foundation, NASA grant NNG04GK98G and NSF grant PHY-0601459 for financial support. RP's work was supported by the Max-Planck Society. MV was supported by LISA Mission Science Office and by JPL's Human Resources Development Fund. LW's work was supported by the Alexander von Humboldt Foundation's Sofja Kovalevskaja Programme, funded by the German Federal Ministry of Education and Research.

References

- [1] Bender P, Danzmann P and the LISA Study Team 1998 "Laser Interferometer Space Antenna for the Detection of Gravitational Waves, Pre-Phase A Report" **MPQ 233** (Garching: Max-Planck-Institut für Quantenoptik)
- [2] Arnaud K A et al. (the Mock LISA Data Challenge Task Force) 2006 *Laser Interferometer Space Antenna: 6th International LISA Symp. (Greenbelt, MD, 19–23 Jun 2006)* ed Merkwowitz S M and Livas J C (Melville, NY: AIP) p 619; *ibid.* p 625
- [3] Arnaud K A et al. (the Mock LISA Data Challenge Task Force and Challenge 1 participants) 2007 *Class. Quant. Grav.* **24** S529
- [4] Arnaud K A et al. (the Mock LISA Data Challenge Task Force) 2007 *Class. Quant. Grav.* **24** S551
- [5] P. Jaranowski, A. Królak, and B. F. Schutz, *Phys. Rev. D* **58**, 063001 (1998).
- [6] A. Królak, M. Tinto, and M. Vallisneri, *Phys. Rev. D*, **70**, 022003 (2004).
- [7] Various authors 2008 Proc. 7th Amaldi Conf. on Gravitational Waves (Sydney, 8–14 July 2007) *J. Phys. Conf. Ser.* **XX**
- [8] Brown D A, Crowder J, Cutler C, Mandel I and Vallisneri M 2007 *Class. Quant. Grav.* **24** S595
- [9] Vallisneri M, Crowder J and Tinto M 2007 "Sensitivity and parameter-estimation precision for alternate LISA configurations" *Preprint* arxiv.org/0710.4369
- [10] Gair J R, Mandel I and Wen L 2007 Proc. 7th Amaldi Conf. on Gravitational Waves (Sydney, 8–14 July 2007), submitted. *Preprint* arXiv:0710.5250
- [11] Prix R and Whelan J T 2007 *Class. Quant. Grav.* **24** S565; *Poster* www.ligo.caltech.edu/docs/G/G070462-00.pdf