

A 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 Cutler⁵, Shane L. Larson⁶, Tyson B. Littenberg⁴, Edward K. Porter¹, Michele Vallisneri⁵, Alberto Vecchio^{7,8} and *the Challenge-2 participants:* Gerard Auger⁹, Leor Barack¹⁰, Arkadiusz Blaut¹¹, Ed Bloomer¹², Duncan A. Brown^{13,14}, Nelson Christensen¹⁵, James Clark¹², Stephen Fairhurst^{13,14}, Jonathan R. Gair¹⁶, Hubert Halloin⁹, Martin Hendry¹², Arturo Jimenez³, Andrzej Królak¹¹, Ilya Mandel¹⁴, Chris Messenger¹², Renate Meyer¹⁷, Soumya Mohanty³, Rajesh Nayak³, Antoine Petiteau⁹, Matt Pitkin¹², Eric Plagnol⁹, Reinhard Prix¹, Emma L. Robinson⁷, Christian Roever¹⁷, Pavlin Savov¹⁴, Alexander Stroeer^{7,8}, Jennifer Toher¹², John Veitch⁷, Jean-Yves Vinet¹⁸, Linqing Wen¹, John T. Whelan¹, Graham Woan¹²

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

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

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

⁴ Department of Physics, Montana State University, Bozeman, MT 59717, USA

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

⁶ Department of Physics, Weber State University, 2508 University Circle, Ogden, UT 84408, USA

⁷ School of Physics and Astronomy, University of Birmingham, Edgbaston, Birmingham B152TT, UK

⁸ Department of Physics and Astronomy, Northwestern University, Evanston, IL, USA

⁹ APC, UMR 7164, University Paris 7 Denis Diderot, 10, rue Alice Domon et Leonie Duquet, 75025 Paris Cedex 13, France

¹⁰ School of Mathematics, University of Southampton, Southampton, SO171BJ, UK

¹¹ Institute of Mathematics, Polish Academy of Sciences, Warsaw, Poland

¹² Department of Physics and Astronomy, University of Glasgow, Glasgow, UK

¹³ LIGO Laboratory, California Institute of Technology, Pasadena, CA 91125

¹⁴ Theoretical Astrophysics, California Institute of Technology, Pasadena, CA 91125

¹⁵ Physics and Astronomy, Carleton College, Northfield, MN, USA

¹⁶ Institute of Astronomy, University of Cambridge, Cambridge, CB30HA, UK

¹⁷ Department of Statistics, The University of Auckland, Auckland, New Zealand

¹⁸ ARTEMIS, Observatoire de la Cote d'Azur-C.N.R.S., 06304 Nice, France

E-mail: Michele.Vallisneri@jpl.nasa.gov

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, which 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/ESA space mission to detect gravitational waves in the 10^{-5} – 10^{-1} Hz range [1], will produce time series consisting of the superposition of the gravitational waves (GWs) from all sources within range, typically thousands. Some of the gravitational 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 similar signals, overlapping in frequency space. Thus, data analysis is integral to the LISA measurement concept, because no physical 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 a computational infrastructure for LISA data analysis, which includes common data formats, standard models of the LISA orbits, noises and measurement, descriptions of the waveforms and populations of targeted GW sources, and so on. 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 datasets with 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 Refs. YYY and ZZZ), 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 moderately 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 participated, 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 parameter probabilities. More important, Challenge 1 helped set the playing field and the computational tools for the more realistic Challenge 2.

Challenge 2, issued in Jan 2007 with results due at this conference, raised the bar by proposing data set 2.1, containing a full population of Galactic binary systems (about 26 million sources); data set 2.2, containing a different realization of this Galaxy, plus 4–6 MBH binary inspirals with single-interferometer signal-to-noise ratios (SNRs) between 10 and 2000 and different coalescence times, and 5 EMRIs with SNRs between 30 and 100; and five data sets (denoted 1.3.1–5, since they were actually released at the time of Challenge 1) of a single EMRI over instrument noise alone. See Ref. WWW for more details about the signal models and the exact content of the data sets. Thirteen collaborations participated (comprising all the researchers listed as participants in the byline, and most task force members), submitting a total of 22 entries. Altogether, Challenge 2 successfully demonstrated the identification of $\sim 20,000$ Galactic binaries, the accurate parameter estimation of MBH inspirals, and the positive detection of EMRIs. In the rest of this paper, we describe some Challenge-2 highlights. The full set of responses, together with write-ups by all participating groups, can be found at the URL XXX; some groups are also contributing proceedings from this conference to JPCS (Ref. YYY).

2. Data sets 2.1 and 2.2: The Galaxy

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

GLIG A collaboration of research groups at institutions in the UK, United States and New Zealand developed a Reversible-Jump Markov Chain Monte Carlo (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 used \mathcal{F} -statistic, template-bank-based matched filtering [2, 3], and submitted parameters for 404 sources in data set 2.1.

MTJPL The Montana State–JPL collaboration used an MCMC code that ran separately for overlapping frequency bands and for different hypothesized numbers of sources; model comparison was then used to choose the most probable number of sources in each band. The collaboration submitted parameter sets for 19,324 source in data set 2.1, and 18,461 in data set 2.2.

PrixWhelanAEI Prix and Whelan used \mathcal{F} -statistic, template-bank-based matched filtering in a hierarchical scheme that enforced trigger coincidence between noise-orthogonal TDI observables. They submitted parameter sets for 1777 sources in data set 2.1, and 1737 sources in 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 a problem of principle: the performance metric that makes the most physical sense would be a statement about how many of the true sources were recovered, and with what parameter errors. For this, however, the reported sources must first be matched one-to-one with the true sources, a nontrivial operation because of the very nature of confusion noise (consider for instance the phenomena of *source blending*, where multiple true sources end up

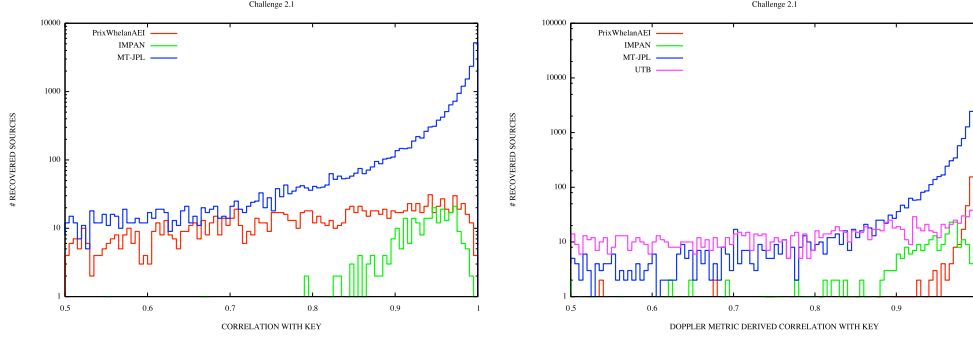


Figure 1. Correlation analysis of Challenge-2.1 Galactic-binary catalogs. Left panel: associated by correlation; right panel: associated by Doppler metric.

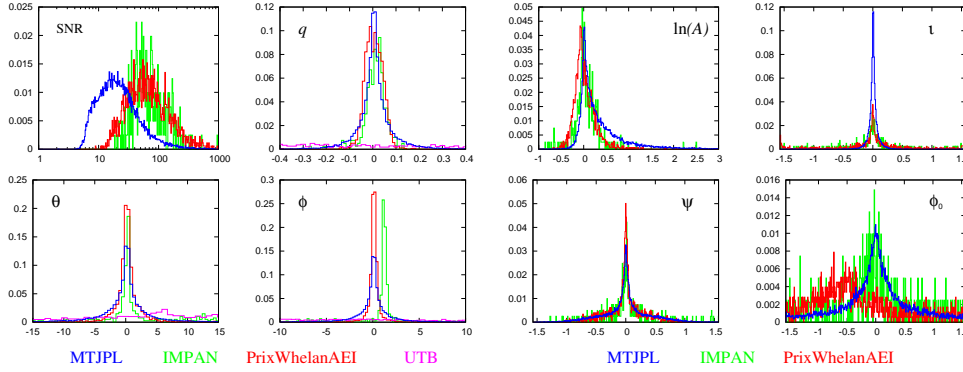


Figure 2. Recovered SNRs and intrinsic and extrinsic parameter errors for Challenge-2.1 Galactic-binary catalogs. True sources and templates are associated by correlation, except for the UTB catalog, for which they are associated by Doppler metric.

being represented by a single search template, and of *template blending*, where multiple templates end up representing a single strong source).

One way to proceed is to associate the reported and true sources that have the strongest signal correlation (limiting the search to the “bright” $\text{SNR} > 2$ true sources that could in principle have been found) [[can true sources be “reused” in this association?]]: in the left panel of figure 1 we show the distribution of correlations generated in this process for data set 2.1. Another way is to associate reported and true 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 (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

Table 1. Recovered SNR and parameter errors for MBH-1. All parameters are defined as in tables 2 and 4 of [?], 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	–	–	–	–	–	–

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; the errors in sky position are within a few degrees; by contrast the errors in the amplitude and in the (extrinsic) orientation angles are larger. The ϕ_0 graph for the PrixWhelanAEI suggests a systematic error in the definition of initial phase. [[Was there some correspondence about this?]]

These challenges demonstrated a solid capability in analyzing signals from the Galaxy and resolving a large number of binaries; 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, targeted work.

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

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

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

MTAEI. Cornish and Porter used MCMC matched filtering in a *frequency-annealed* scheme where shorter, lower-frequency templates are used in the initial phases of the search and 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;

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 ~ 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 (complete results can be found at the URL XXX), together with the SNR recovered by the best-fit candidate of each group, 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 2. Recovered SNR and parameter errors for MBH-4, given as in table 1. The true (optimal) SNR is **174.12**.

	SNR	$\Delta M_c/M_c$ $\times 10^{-6}$	$\Delta \mu/\mu$ $\times 10^{-4}$	$\Delta t_c/t_c$ $\times 10^{-6}$	$\Delta \beta$ $\times 10^{-2}$	$\Delta \lambda$ $\times 10^{-2}$	$\Delta D/D$ $\times 10^{-3}$	$\Delta \iota$ $\times 10^{-3}$	$\Delta \psi$ $\times 10^{-3}$	$\Delta \phi_c$ $\times 10^{-1}$
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., VCT). In table 1, we see that compared to MTAEI, the JPLCT search for MBH-1 locked onto a secondary probability maximum with only slightly lower SNR_{best} , but with sky positions off by several degrees, which also led to errors in the other parameters (the JPLCT authors think this was caused by first subtracting off a rough match for MBH-1, then subtracting off the resolvable Galactic binaries, and then refining the MBH search). MBH-4 (table 1) is an interesting example of a true (at least) bimodal probability distribution. MTAEI and JPLCT reported two candidates at rather different sky locations, with relative probability ratios of 1:1 and 1.18:1, respectively. The position and orientation of this source relative to LISA, which resulted in a very weak signal in one of the noise-orthogonal observables, probably concurred in degrading LISA’s sensitivity to sky position.

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

4. Data sets 1.3.X: EMRIs

Three groups reported parameters for the EMRIs in the 1.3.1–1.3.4 data sets. No group tackled the problem of detecting these systems over the Galactic background in data set 2.2.

BGGP Babak and colleagues used MCMC matched filtering, introducing progressively longer templates (a *time-annealed* scheme);

EtfAG Gair, Mandel and Wen used a TF track search insensitive to amplitude and orientation parameters;

MT Cornish used MCMC matched filtering on 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 secondary probability maxima (indeed, several of them 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 an

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 [?]. 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

order of magnitude more precise in determining the intrinsic parameters. 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 ??, but the prospects for accurate parameter estimation are still uncertain, and a good focus for further challenges.

5. Conclusion

[Still to be written.]

We're happy:

- Participation is strong (but we need more, try 1B!)
- A lot of work is being accomplished
- Results are comforting
- Were showing that LISA data analysis is possible
- Difficulty is incremental, but no big issues of principle are foreseen
- The MLDC infrastructure (LISAtools) can be used for many other experiments outside the mainline challenges

Acknowledgments

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. MV was supported by LISA Mission Science Office and by JPL's Human Resources Development Fund.

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] P. Jaranowski, A. Królak, and B. F. Schutz, Phys. Rev. D **58**, 063001 (1998).
- [3] A. Królak, M. Tinto, and M. Vallisneri, Phys. Rev. D, **70**, 022003 (2004).