Report on the first round of the Mock LISA Data Challenges

K A Arnaud[1], **G Auger**[2], **S Babak**[3], **J G Baker**[1], **M J Benacquista**[4],
E Bloomer[5], **D A Brown**[6,7], **J B Camp**[8], **J K Cannizzo**[8], **N Christensen**[9],
J Clark[5], **N J Cornish**[10], **J Crowder**[10,11], **C Cutler**[7,11], **L S Finn**[12],
H Halloin[2], **K Hayama**[4], **M Hendry**[5], **O Jeannin**[2], **A Królak**[13],
S L Larson[14], **I Mandel**[6], **C Messenger**[5], **S Mohanty**[4],
R Nayak[4], **K Numata**[8], **A Petiteau**[2], **M Pitkin**[5], **E Plagnol**[2],
E K Porter[3,10], **R Prix**[3], **C Roever**[15], **A Stroeer**[16,17],
R Thirumalainambi[18], **D E Thompson**[18], **J Toher**[5], **R Umstaetter**[15],
M Vallisneri[7,11], **A Vecchio**[16,17], **J Veitch**[16], **J-Y Vinet**[19], **J T Whelan**[3] and
G Woan[5]

[1] Gravitational Astrophysics Laboratory, NASA Goddard Space Flight Center,
8800 Greenbelt Rd, Greenbelt, MD 20771, USA
[2] AstroParticule et Cosmologie (APC), UMR 7164, Université Paris 7 Denis Diderot,
10 rue Alice Domon et Léonie Duquet, F-75205 Paris Cedex 13, France
[3] Max-Planck-Institut für Gravitationsphysik (Albert-Einstein-Institut), Am Mühlenberg 1,
D-14476 Golm bei Potsdam, Germany
[4] Center for Gravitational Wave Astronomy, University of Texas at Brownsville, Brownsville,
TX 78520, USA
[5] Department of Physics and Astronomy, University of Glasgow, Glasgow G12 8QQ, UK
[6] LIGO Laboratory, California Institute of Technology, Pasadena, CA 91125, USA
[7] Theoretical Astrophysics, California Institute of Technology, Pasadena, CA 91125, USA
[8] Laboratory for Gravitational Physics, Goddard Space Flight Center, Greenbelt, MD 20771, USA
[9] Physics and Astronomy, Carleton College, Northfield, MN, USA
[10] Department of Physics, Montana State University, Bozeman, MT 59717, USA
[11] Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA
[12] Center for Gravitational Wave Physics, The Pennsylvania State University, University Park,
PA 16802, USA
[13] Institute of Mathematics, Polish Academy of Science, Warsaw, Poland
[14] Department of Physics, Weber State University, 2508 University of Circle, Ogden, UT 84408,
USA
[15] Department of Statistics, The University of Auckland, Auckland, New Zealand
[16] School of Physics and Astronomy, University of Birmingham, Edgbaston, Birmingham
B15 2TT, UK
[17] Department of Physics and Astronomy, Northwestern University, Evanston, IL 60208, USA
[18] NASA Ames Research Center, Moffett Field, CA, USA
[19] Department ARTEMIS, Observatoire de la Côte d'Azur, BP 429, 06304 Nice, France

E-mail: benacquista@phys.utb.edu

Abstract
The Mock LISA Data Challenges (MLDCs) have the dual purpose of fostering the development of LISA data analysis tools and capabilities, and demonstrating the technical readiness already achieved by the gravitational-wave community in distilling a rich science payoff from the LISA data output. The first round of MLDCs has just been completed: nine challenges consisting of data sets containing simulated gravitational-wave signals produced either by galactic binaries or massive black hole binaries embedded in simulated LISA instrumental noise were released in June 2006 with deadline for submission of results at the beginning of December 2006. Ten groups have participated in this first round of challenges. All of the challenges had at least one entry which successfully characterized the signal to better than 95% when assessed via a correlation with phasing ambiguities accounted for. Here, we describe the challenges, summarize the results and provide a first critical assessment of the entries.

1. Introduction

At the LISA International Science Team (LIST) meeting of December 2005, the Working Group on Data Analysis (LIST-WG1B) decided to organize several rounds of MLDCs with the dual purposes of (i) fostering the development of LISA data analysis tools and capabilities and (ii) determining the technical readiness already achieved by the gravitational-wave community for distilling a rich science payoff from the LISA data output. These challenges are meant to be blind tests, but not contests. The intent is to encourage the quantitative comparison of results, analysis methods and implementations.

A MLDC Task Force was constituted at the beginning of 2006 and has been working since then to formulate challenge problems, develop standard models of the LISA mission and gravitational-wave (GW) sources, provide computing tools (e.g. LISA response simulators and source waveform generators), establish criteria for the evaluation of the analyses and provide any technical support necessary to the challenge participants. The first round of challenges involve the distribution of several data sets, encoded in a simple standard format, and containing combinations of realistic simulated LISA noise with the signals from one or more GW sources with parameters which were unknown to the participants. The participants were then asked to return the maximum amount of information about the sources and to produce technical notes detailing their work.

The release of the first round of challenge data sets was announced in June 2006 at the Sixth LISA International Symposium hosted by the Goddard Space Flight Center in Greenbelt, Maryland [1, 2]. John Baker (a member of the MLDC Task Force who did not participate in the first round) was appointed as MLDC1 coordinator. The coordinator was responsible for generating the challenge data sets, receiving the results from the participants, and posting both the key data files and results as soon as possible after the submission deadline of 4 December 2006.

The challenge data sets include a total of 9 years long data sets which are described in detail on the MLDC website [3], the Task Force wiki [4] and the Omnibus document for Challenge 1 [5]. The challenge data sets are broadly grouped into three categories: (1.1) white

Table 1. Groups that participated in the Mock LISA Data Challenge 1. The challenges for which each group submitted results are marked by (●).

| | Galactic binaries | | | | | | | Massive | |
| | Single source | | | Multiple sources | | | | Black holes | |
Group	1.1.1a	1.1.1b	1.1.1c	1.1.2	1.1.3	1.1.4	1.1.5	1.2.1	1.2.2
AEI	●	●	●	●	●	●	●		
Ames	●	●		●	●				
APC	●								
Goddard								●	
GLIG			●						
Królak	●	●	●		●				
JPL/Caltech								●	
MT/AEI								●	●
MT/JPL	●	●	●	●	●	●	●		
UTB	●	●	●	●	●	●			

dwarf binaries (WDs), (1.2) supermassive black holes (SMBHs) and (1.3) extreme mass ratio inspirals (EMRIs). The problem of detection of EMRIs is considered more difficult than the others [6], so the deadline for submission of results for the 1.3 challenges is extended to June 2007. Consequently, in this paper we will discuss the results of Challenges 1.1 and 1.2.

The WD challenges consist of three single source data sets with the GW frequency around 1 mHz (1.1.1a), 3 mHz (1.1.1b) and 10 mHz (1.1.1c) and four multiple source data sets with isolated sources of known (1.1.2) and unknown (1.1.3) sky locations and frequencies, and overlapping sources with a low (1.1.4) and high (1.1.5) density of sources (0.095 and 0.35, respectively) in frequency space. The SMBH challenges consist of two single source data sets. In one (1.2.1) the SMBH binary merges during the observation time, and in the other (1.2.2) the merger takes place between one and three months after the end of the data set. The source parameters were chosen with a balance between realism and computational efficiency as described in the announcement of the release [1].

2. Overview of MLDC1 submissions

Ten groups submitted results for MLDC1 by the deadline. These results have been posted on the MLDC website. They include the technical notes submitted by the challenge participants and the files with the 'best parameter fits' for the data sets. Table 1 provides a summary of the groups and their submissions for MLDC1.

With the exception of Challenges 1.1.5 and 1.2.2, every challenge data set was analysed by at least three groups. Here, we briefly summarize the approaches used by each group. More detailed descriptions from many of the groups can be found elsewhere in these proceedings [7–17] or in the technical notes on the MLDC web page [3]. Several groups used variations on matched filtering methods on many of the challenges. The Ames group at the NASA Ames Research Center employed a user-refined grid search on a number of the WD challenges. The AEI group and Andrzej Królak both used grid-based methods. The Global LISA Inference Group (GLIG), Montana-JPL and Montana-AEI groups employed variations on Markov Chain Monte Carlo (MCMC) methods [18]. The JPL-Caltech group used a multi-stage approach that combined time–frequency methods with grid-based and MCMC searches. The Montana-JPL group also used a genetic algorithm [19]. The APC group has also implemented a hierarchical

approach which first matches the annual amplitude modulation and then follows with a full-matched filtering. Two groups did not use matched filtering at all. The UTB group used a tomographic search that employed the Radon transform [20] while the Goddard group at the NASA Goddard Space Flight Center developed a time–frequency method that uses the Hilbert–Huang transform [21]. Although some of these groups have well-developed and mature algorithms, most groups are currently in various stages of development and so many of the entries are incomplete or suffer from known bugs which could not be hunted down before the December deadline. Some of the algorithms that are under development are meant to be part of a hierarchical search and so they only return a subset of the parameters needed to fully characterize the source.

3. Assessment

We evaluate the performance of all entries using the noise orthogonal pseudo A and E channels [22]:

$$A = (2X - Y - Z)/3, \qquad E = (Z - Y)/\sqrt{3}, \tag{1}$$

where X, Y and Z are the standard TDI variables. The wide variety of approaches and maturity of the algorithms makes it difficult to develop a single assessment that can adequately compare all entries. For those entries that have returned enough parameters to sufficiently generate a recovered waveform, we can compare the recovered waveform h_{rec}, with the waveform generated from the 'true' parameters h_{key} using

$$\Delta \chi^2 = \frac{(h_{key} - h_{rec} | h_{key} - h_{rec})}{D} \tag{2}$$

where $(*|*)$ is the noise (S_n) weighted inner product summed over channels $(i = A, E)$, defined by

$$(a|b) = 2 \int_{f_{min}}^{f_{max}} \frac{\sum_{i=A,E} (\tilde{a}_i b_i + a_i \tilde{b}_i)}{S_n} \, df \tag{3}$$

and D is the dimension of the parameter space used to generate the templates. We realize that $\Delta \chi^2$ is not a perfect figure of merit as, for example, it does not account for deduced uncertainties in the recovered parameters. It is however easy to compute and is quite sufficient to indicate whether the recovered parameters differ greatly from those used for the key waveform. We can also compute the signal-to-noise ratio (SNR) for both h_{key} and h_{rec} using

$$SNR = \frac{(s|h)}{\sqrt{(h|h)}} \tag{4}$$

and compare the recovered SNR with the key SNR. Finally, we calculate the correlation between h_{key} and h_{rec} with

$$C = \frac{(h_{key} | h_{rec})}{\sqrt{(h_{key} | h_{key})(h_{rec} | h_{rec})}}. \tag{5}$$

The $\Delta \chi^2$ measure is sensitive to the difference in amplitude between h_{key} and h_{rec} while the SNR and correlation are both insensitive to the amplitude of h_{rec}.

Some groups reported a known ambiguity in the initial phase and polarization angles with results being given modulo $\pi/2$ in the polarization angle and modulo π in the initial phase [23]. The results would be indistinguishable under $\Delta \chi^2$, SNR, or C tests, but would have been apparent in the parameter estimations errors. We have accounted for this ambiguity by choosing the values of ψ and ϕ_0 that minimize the error in ψ. If there are additional errors in

Table 2. The performance of challenge entries on the single binary challenges as calculated using $\Delta\chi^2$, SNR and C. The correction of the initial phase by a factor of π or $\pi/2$ is indicated by an asterisk (*).

Group	$\Delta\chi^2$	SNR	C
Challenge 1.1.1a ($\mathrm{SNR_{key}} = 51.137$)			
AEI	8.095	50.604	0.989
Ames	7.155	51.032	0.997
APC	423.406	−8.007	−0.135
APC*	229.115	50.385	0.990
Królak	778.888	0.933	−0.004
Królak*	1.036	51.038	0.999
MT/JPL (BAM)	1.902	51.178	0.998
MT/JPL (GA)	1.796	51.138	0.998
Challenge 1.1.1b ($\mathrm{SNR_{key}} = 37.251$)			
AEI	47.913	33.104	0.874
Ames	64.371	32.067	0.822
Królak	841.074	−37.038	−0.996
Królak*	2.566	37.038	0.996
MT/JPL (BAM)	7.735	36.856	0.980
MT/JPL (GA)	8.371	36.808	0.979
Challenge 1.1.1c ($\mathrm{SNR_{key}} = 101.390$)			
AEI	2399.307	−14.373	−0.144
GLIG	1788.991	14.496	0.142
Królak	5997.595	−98.126	−0.968
Królak*	97.603	98.126	0.968
MT/JPL (BAM)	945.541	63.383	0.623
MT/JPL (GA)	1376.143	43.564	0.424

the returned value of ϕ_0, they would remain after correcting for this ambiguity. These errors can arise from using different conventions for generating the waveform from the physical parameters or they can arise from determining the parameters in the form of a trigonometric function and then choosing the wrong quadrant when applying the inverse operation to obtain the angle. Obviously, a difference of π in the initial phase can significantly degrade the performance of an entry as calculated using $\Delta\chi^2$, SNR or C. Consequently, we have also computed these measures with the initial phase shifted by π if necessary.

Another measure of the success of a given algorithm is the accuracy with which it returns specific parameters. This approach allows us to also evaluate those entries which do not return enough parameters to generate h_{rec}. For each parameter λ_i, we can determine the difference between the key parameter and the recovered parameter using

$$\Delta\lambda = \lambda_{\mathrm{key}} - \lambda_{\mathrm{rec}}. \tag{6}$$

We note that it is not necessarily appropriate to use the Fisher information matrix (FIM) to determine the quality of parameter recovery. If the algorithm settles on a secondary maximum of the likelihood function or there are other systematic errors, then the results can be far from the regime of validity for the FIM approximation to expected errors in parameter estimation. For more details, see Vallisneri's review of the FIM [24].

The white dwarf binary challenges required the recovery of seven parameters to fully characterize each source. These parameters are the amplitude $\mathcal{A}$, the frequency f, the sky location θ, ϕ, the angle of inclination ι, the polarization angle ψ and the initial phase ϕ_0. In table 2, we list the values of the measures for each challenge entry for Challenges 1.1.1—

Table 3. The performance of challenge entries on the single binary challenges as calculated using recovered parameter differences.

Group	Δf (nHz)	$\Delta\theta$	$\Delta\phi$	$\Delta\ln A$	$\Delta\iota$	$\Delta\psi$	$\Delta\phi_0$
			Challenge 1.1.1a				
AEI	−1.208	−0.018	0.001	−0.078	−0.101	0.157	−0.065
Ames	−1.889	−1.159	3.127	0.337	0.503	0.181	−0.126
APC	1.343	−0.030	−0.011	0.807	0.217	0.174	1.202
Królak	0.980	0.028	−0.008	0.113	0.180	0.208	−2.089
MT/JPL (BAM)	−1.367	−0.015	−0.008	−0.046	−0.084	0.196	−0.228
MT/JPL (GA)	−1.044	−0.013	−0.003	−0.077	−0.091	0.224	−0.308
UTB	−3.209	0.143	0.603	–	–	–	–
			Challenge 1.1.1b				
AEI	0.399	−0.049	0.001	−0.009	−0.045	0.020	0.432
Ames	−21.098	−0.606	0.004	0.171	0.048	0.028	2.173
Królak	0.341	0.037	−0.004	−0.112	−0.042	−0.042	−3.098
MT/JPL (BAM)	0.434	−0.040	0.003	−0.025	−0.042	0.029	0.097
MT/JPL (GA)	0.314	−0.039	0.003	−0.044	−0.044	0.030	0.117
UTB	−4.299	0.198	0.007	–	–	–	–
			Challenge 1.1.1c				
AEI	−0.405	0.012	−0.001	0.312	−0.159	0.127	1.501
GLIG	154.850	0.306	0.178	0.341	0.939	0.722	2.413
Królak	−5.210	0.059	−0.010	−0.194	−0.268	0.451	2.747
MT/JPL (BAM)	−0.330	0.008	−0.001	0.309	0.033	−0.609	2.148
MT/JPL (GA)	0.311	0.013	−0.001	0.652	−1.062	−0.614	2.026
UTB	8.577	0.139	0.066	–	–	–	–

with the exception of the UTB entry. Since the UTB algorithm only returns frequency (in intervals of resolvable frequency bins $\sim$32 nHz) and sky position, it cannot be included in this comparison. However, it can be included in the comparison of parameter differences given in table 3. It should be noted that the GLIG cluster crashed before the completion of the algorithm and therefore the MCMC chain did not have the chance to burn into the final values. We also note that although the measures of Challenge 1.1.1c in table 2 seem to be quite bad, most of this is due to the fact that several groups used the low frequency approximation (which ignores the transfer function) in their analyses. Nonetheless, it is important to note that the accuracy in recovery of the sky positions and frequencies is still comparable to Challenges 1.1.1a and 1.1.1b. When corrected for an initial phase error, the entry by Królak (which incorporated the transfer function in the analysis) can be seen to have done quite well.

The multi-source challenges present a different problem for assessment, since there is the possibility of false positives and false negatives. Consider the possibility in which the recovered parameters for one binary out of many are wildly off. If one were to use the correlation between each recovered template and one of the source binaries, it is possible to count the recovered binary as a false positive and the true binary as a false negative. Given the phasing issues that were apparent in Challenges 1.1.1, it is quite likely that there will be several false positive/false negative pairs if such a correlation analysis is used. Consequently, we determine which recovered template goes with which source by looking for template/source pairs that are within one resolvable frequency bin of each other. The overall success of the recovery can be measured using the combined signal of the entire population of recovered binaries as h_{rec} and comparing with the entire population of the true source binaries as h_{key}.
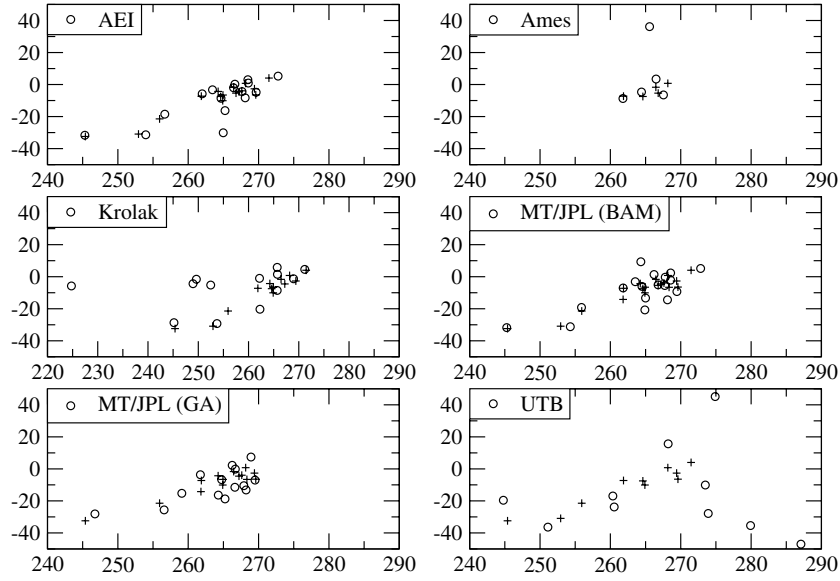
Table 4. The performance of challenge entries on the verification binary Challenge 1.1.2 as calculated using $\Delta\chi^2$, SNR and C. Since every group returned the full 20 binaries, the dimension of the model is $D = 140$. The SNR of the key is 544.952.

Group	$\Delta\chi^2$	SNR	C
AEI	1443.59	339.262	0.624
Ames	227.007	516.471	0.948
MT/JPL (BAM)	19.02	544.165	0.998
MT/JPL (GA)	194.46	519.712	0.954

Table 5. The performance of challenge entries on the isolated binary Challenge 1.1.3 as calculated using $\Delta\chi^2$, SNR and C. Since not all groups returned the full 20 binaries, the dimension of the model is calculated by $D = 7 \times N_{\text{rec}}$, where N_{rec} is the number of recovered binaries. The SNR of the key is 122.864. The correction of the initial phase by a factor of π or $\pi/2$ is indicated by an asterisk (*).

Group	$\Delta\chi^2$	SNR	C	N_{rec}
AEI	67.67	89.789	0.726	16
Ames	48.42	104.236	0.841	13
Królak	204.36	9.558	0.080	20
Królak*	150.95	38.770	0.323	20
MT/JPL (BAM)	44.81	98.838	0.797	19
MT/JPL* (BAM)	4.16	121.64	0.981	19
MT/JPL (GA)	19.66	113.797	0.914	18

In this case, we note that the dimension D of the recovered parameter space depends upon the number of recovered sources.

The Challenge 1.1.2 data set contained 20 'verification' binaries whose frequencies and sky location were given to the participants. Six of these binaries were taken from the list of known binaries available on Gijs Nelemans website [25] and the remaining 14 were simulated binaries. Of the four groups that submitted entries for Challenge 1.1.2, two of them did not use the additional information of sky location and frequency in their searches. The Montana/JPL group used the same search algorithms as they used in all of the 1.1 challenges, and the UTB group used their tomographic algorithm to obtain the frequency and sky location. The UTB group successfully recovered 17 of the 20 source binaries. The other three groups successfully recovered all 20 source binaries, although the AEI group had very low correlation with the two highest frequency binaries. This is understandable since the AEI group used the low frequency approximation (which is valid for frequencies below about 3 mHz) for calculating their templates and the two highest frequency binaries were at frequencies above 6 mHz. These two binaries are also responsible for most of the loss in correlation for this entry. The performance of the three groups that returned the complete parameterization of each binary recovered is given in table 4.

Challenge 1.1.3 also contained 20 binaries isolated in frequency space between 0.5 mHz and 10 mHz. In this challenge, all binaries were drawn from the simulation and all parameters were blind. Five groups submitted entries, although the UTB group again only provided frequency and sky location. The UTB group successfully identified 14 of the 20 binaries. The performance of the other four entries is given in table 5. The sky locations returned by each group are compared with the source positions in figure 1.

Figure 1. Recovered sky positions from each entry for Challenge 1.1.3. The source positions are indicated by + and the recovered positions are indicated by ○. Each plot includes only those sources which are within a frequency bin of a recovered source. The plots are of ecliptic latitude versus longitude in degrees.

Table 6. The performance of challenge entries on the overlapping binary Challenges 1.1.4 and 1.1.5 as calculated using $\Delta\chi^2$, SNR and C. Since not all groups returned the full number of binaries, the dimension of the model is calculated by $D = 7 \times N_{rec}$, where N_{rec} is the number of recovered binaries. N_{match} is the number of recovered binaries that matched in frequency with a binary in the key.

Group	$\Delta\chi^2$	SNR	C	N_{rec}	N_{match}
Challenge 1.1.4 (SNR$_{key}$ = 201.129)					
AEI	85.63	159.893	0.792	26	16
MT/JPL (BAM)	6.19	197.828	0.976	43	39
Challenge 1.1.5 (SNR$_{key}$ = 178.261)					
AEI	519.21	116.822	0.654	5	4
MT/JPL (BAM)	11.96	172.582	0.963	27	23

The remaining two white dwarf challenges each contained approximately 45 overlapping sources. Challenge 1.1.4 contained 45 sources in a 15μ Hz band starting at 3 mHz, while Challenge 1.1.5 contained 33 sources in a ±1.5 μHz band centred on 3 mHz. Challenge 1.1.4 had an average density of 0.095 sources per resolvable frequency bin, and Challenge 1.1.5 had an average density of 0.35 sources per bin. These challenges were actually more difficult than these source densities might indicate since in both cases there were at least three frequency bins which contained at least two binaries each. There was one case in Challenge 1.1.5 with three binaries in one frequency bin. Despite this additional complication, both groups managed to recover a respectable number of sources that matched with binaries in the key. The performance of the two groups that submitted complete parameter sets for each binary recovered is listed in table 6.

Table 7. The performance of challenge entries on the supermassive black hole binary Challenges 1.2.1 and 1.2.2. The results after a correction of the initial phase are indicated by an asterisk (*).

Group	$\Delta\chi^2$	SNR	O_A	O_E	$\max_{\phi_0}(O_X)$
Challenge 1.2.1 ($\mathrm{SNR_{key}} = 667.734$)					
JPL	261.48	664.47	0.994	0.996	0.9955
MT/AEI	10 289.29	524.29	0.790	0.791	0.9998
MT/AEI*	105.50	662.87	0.998	0.998	0.9998
Challenge 1.2.2 ($\mathrm{SNR_{key}} = 104.19$)					
MT/AEI	1.41	104.29	0.997	0.998	0.9955

The supermassive black hole challenges required recovery of nine parameters describing the source: the chirp mass $\mathcal{M}$, the reduced mass μ, the luminosity distance D_L, the time of coalescence t_c, the sky location θ, ϕ, the initial angle of inclination ι, the initial polarization angle ψ and the initial orbital phase ϕ_0. As with the white dwarf challenges, the quality of the recovered signal can be described by the measures given in equations (2) and (4). In place of the correlation, we compute several overlaps using

$$O_\alpha = \frac{\left(h^\alpha_{\mathrm{key}} \big| h^\alpha_{\mathrm{rec}}\right)}{\sqrt{\left(h^\alpha_{\mathrm{key}} \big| h^\alpha_{\mathrm{key}}\right)\left(h^\alpha_{\mathrm{rec}} \big| h^\alpha_{\mathrm{rec}}\right)}} \tag{7}$$

where α denotes the particular TDI channel being used. In order to mitigate the effects of a possible error in the initial phase, we have also computed O_X, maximized over the phase:

$$\max_{\phi_0}(O_X) = \sqrt{\left(h^X_{\mathrm{rec}} \big| h^X_{\mathrm{key}}(\phi_0 = 0)\right)^2 + \left(h^X_{\mathrm{rec}} \big| h^X_{\mathrm{key}}(\phi_0 = \pi/2)\right)^2}. \tag{8}$$

There were two groups that returned a full characterization of the signal for Challenge 1.2.1. The Montana/AEI group had a constant phase difference, and when this phase is corrected the performance of both the JPL/Caltech and Montana/AEI groups is quite good. The Goddard group is developing a new algorithm using the Hilbert–Huang transform that is in a very preliminary stage and has only returned $\mathcal{M}$ and t_c for this challenge. Because of a known secondary maximum in sky location, we also check the antipodal sky position:

$$\theta \to -\theta, \qquad \phi \to \phi \pm \pi. \tag{9}$$

However, this adjustment also requires a change in the values of the inclination and polarization angles as well. This is accomplished by substituting the initial returned values of θ, ϕ, ι and ψ into

$$\cos\iota = \cos\theta \sin\Theta \cos(\phi - \Phi) + \cos\Theta \sin\theta \tag{10}$$

$$\tan\psi = \frac{\sin\theta \cos(\phi - \Phi) \sin\Theta - \cos\Theta \cos\theta}{\sin\Theta \sin(\phi - \Phi)}, \tag{11}$$

and solve these equations for the orientation angles of the orbital angular momentum vector, (Θ, Φ). Once we have the values of (Θ, Φ), we then use these values and the antipodal sky position from equations (9) in the above equations (10) and (11) to determine the new values of ι and ψ. Once these have been found, the new values can be used to provide a more realistic estimate of the error in the returned values. We have applied this transformation to the JPL/Caltech entry and also adjusted the polarization phase for the Montana/AEI entry in Challenge 1.2.1. Only one group (Montana/AEI) submitted an entry for Challenge 1.2.2. The measures for each submission under both of these challenges are given in table 7. We have also determined the errors in the recovered parameters for all entries in Challenges 1.2.1 and 1.2.2. These are presented in table 8.

Table 8. The performance of challenge entries on the supermassive black hole binary Challenges 1.2.1 and 1.2.2 on the estimation of recovered parameters. The angles are absolute errors, all other errors are relative.

Group	$\Delta\mathcal{M}/\mathcal{M}$ $(\times 10^{-4})$	$\Delta\mu/\mu$	$\Delta D_L/D_L$	$\Delta t_c/t_c$ $(\times 10^{-6})$	$\Delta\theta$	$\Delta\phi$	$\Delta\iota$	$\Delta\psi$	$\Delta\phi_0$
				Challenge 1.2.1 (reported values)					
JPL	7.35	0.011	1.101	3.35	1.030	−3.170	1.32	−2.65	0.004
MT/AEI	0.98	0.001	0.042	0.26	0.001	0.001	0.02	3.14	0.004
Goddard	434.00	–	–	113.00	–	–	–	–	–
				Challenge 1.2.1 (angle adjusted values)					
JPL	7.35	0.011	1.101	3.35	−0.043	−0.032	−0.58	−0.31	0.004
MT/AEI	0.98	0.001	0.042	0.26	0.001	0.001	0.02	−0.00	0.004
				Challenge 1.2.2					
MT/AEI	3.09	0.037	0.273	182.00	0.019	0.005	−0.71	−2.16	−0.002

4. Conclusions

The first round of the Mock LISA Data Challenges successfully attracted over ten groups to work on the problem of LISA data analysis. These groups attacked several of the challenges with a variety of different approaches. The algorithms and codes used in the challenges were at different levels of maturity and completeness of the pipelines. Nonetheless, all challenges had at least one entry which successfully characterized the signal to better than 95% when assessed via the correlation with phasing ambiguities accounted for. In the overlapping source challenge 1.1.5, one group was able to recover true binaries at a source density of ∼0.25. Most groups also discovered small bugs or discrepancies in definitions of some of the parameters used to characterize the signal. The first round entries were a success, especially considering that most groups had less than five months to work out interfacing issues and assess their codes. Those groups that participated in round 1 have begun implementing the lessons learned for use in round 2 [26], which is another successful outcome of the challenges. The MLDC Task Force has also begun to address the issues that have been raised as regards assessment of the entries. As the assessments become more refined in future challenges, we anticipate developing assessments for the true LISA data for which there is no key file.

Acknowledgments

References

[1] Arnaud K A *et al* 2006 *AIP Conf. Proc.* **873** 619 (*Preprint* gr-qc/0609105)
[2] Arnaud K A *et al* 2006 *AIP Conf. Proc.* **873** 625 (*Preprint* gr-qc/0609106)
[3] Mock LISA Data Challenge Homepage astrogravs.nasa.gov/docs/mldc
[4] Mock LISA Data Challenge Task Force wiki www.tapir.caltech.edu/dokuwiki/listwg1b:home
[5] Mock LISA Data Challenge Task Force, 'Document for Challenge 1', svn.sourceforge.net/viewvc/lisatools/Docs/challenge1.pdf
[6] Gair J R, Barack L, Creighton T, Cutler C, Larson S L, Phinney E S and Vallisneri M 2004 *Class. Quantum Grav.* **21** 1595
[7] Auger G *et al* 2007 *Class. Quantum Grav.* submitted
[8] Camp J B, Cannizzo J K and Numata K 2007 *Phys. Rev.* D **75** 061101 (*Preprint* gr-qc/0701148)
[9] Crowder J and Cornish N 2007 *Class. Quantum Grav.* **24** S575
[10] Brown D *et al* 2007 *Class. Quantum Grav.* **24** S595
[11] Królak A 2007 *Class. Quantum Grav.* submitted
[12] Nayak R K, Mohanty S D and Hayama K 2007 *Class. Quantum Grav.* **24** S587
[13] Porter E and Cornish N 2007 *Class. Quantum Grav.* **24** S501
[14] Prix R and Whelan J T 2007 *Class. Quantum Grav.* **24** S565
[15] Roever C *et al* 2007 *Class. Quantum Grav.* **24** S521
[16] Stroeer A *et al* 2007 *Class. Quantum Grav.* **24** S541
[17] Thompson D E and Thirumalainambi R 2007 *Class. Quantum Grav.* submitted
[18] Gilks W R, Richardson S and Spiegelhalter D J 1996 *Markov Chain Monte Carlo in Practice* (Boca Raton, FL: Chapman and Hall/CRC Press)
[19] Holland J 1975 *Adaptation in Natural and Artificial Systems* (Ann Arbor, MI: University of Michigan Press)
[20] Deans S R 1983 *The Radon Transform and Some of its Applications* (New York: Wiley)
[21] Huang N *et al* 1998 *Proc. R. Soc. Lond.* **454** 903
[22] Prince T A, Tinto M, Larson S L and Armstrong J W 2002 *Phys. Rev.* D **66** 122002
[23] Cornish N J and Larson S L 2003 *Phys. Rev.* D **67** 103001
[24] Vallisneri M 2007 Use and abuse of the Fisher information matrix in the assessment of gravitational-wave parameter-estimation prospects *Preprint* gr-qc/0703086
[25] Nelemans G LISA Wiki www.astro.kun.nl/~nelemans/dokuwiki
[26] Arnaud K A *et al* 2007 *Class. Quantum Grav.* **24** S551