Abstract. The LISA Pathfinder data analysis team has been developing in the last years the infrastructure and methods required to run the mission during flight operations. These are gathered in the LTPDA toolbox, an object oriented MATLAB toolbox that allows all the data analysis functionalities for the mission, while storing the history of all operations performed to the data, thus easing traceability and reproducibility of the analysis. The parameter estimation methods in the toolbox have been applied recently to data sets generated with the OSE (Off-line Simulations Environment), a detailed LISA Pathfinder non-linear simulator that will serve as a reference simulator during mission operations. These simulations, so called operational exercises, are the last verification step before translating these experiments into tele-command sequences for the spacecraft, producing therefore very relevant datasets to test our data analysis methods. In this contribution we report the results obtained with three different parameter estimation methods during one of these operational exercises.

1. Introduction
The main objective of the LISA Pathfinder (LPF) mission [1 2] is to characterise the purity of the free fall of two test masses in nominal geodesic motion in its orbit around the Lagrange point L1. The main scientific objective of the mission is specified in a spectral density of relative acceleration fluctuations between test masses of $S_{\Delta a} \leq 3 \times 10^{-14} \text{m s}^{-2} \text{Hz}^{-1/2}$ at 1 mHz. Such an unprecedented metrology sensitivity in a space mission would be considered as a technology achievement sound enough to guarantee a successful development of a future gravitational wave observatory in space. The most mature concept for this space gravitational wave mission is LISA [3], which was recently abandoned as a joint ESA/NASA effort, although the concept is still being actively pursued independently by both space agencies.
An important aspect of the LISA Pathfinder mission is the need of a flexible data analysis infrastructure able to help the taking of decisions since, during operations, the LPF must be understood as an in-flight experiment with hundreds of channels being sampled and a tight schedule of experiments to be executed sequentially.

The LISA Pathfinder data analysis team has been developing in the last years the infrastructure and methods required to run the mission during flight operations. These are gathered in the LTPDA toolbox [4], an object oriented MATLAB toolbox that allows all the data analysis functionalities for the mission, while storing the history of all operations performed to the data, thus easing traceability and reproducibility of the analysis.

In this contribution we focus on three parameter estimation methods—time-domain linear estimation, time-domain non-linear estimation and Markov Chain Monte Carlo in frequency domain—implemented in the toolbox and the comparison of their results when applied with simulated data. The aim of the team when performing this exercise was twofold: to compare the consistency between methods and to test themselves against the most realistic data available.

2. The OSE simulator and the sixth operational exercise

The OSE (Off-line Simulations Environment) is a detailed LISA Pathfinder non-linear simulator that will serve as a reference simulator during mission operations. This simulator is the one being used during a series of runs being regularly executed, the so called operational exercises. These exercises with simulated data aim at testing the on-orbit experiments in a realistic environment in terms of software and time constraints. The operational exercises are the last verification step before translating experiments into tele-command sequences for the spacecraft, producing therefore very relevant datasets to test our data analysis methods.

The data being analysed here correspond to the sixth operational exercise, which was run on April 2011. The simulated data correspond to two runs—called investigations—of around 8 hours duration. In each of these a sequence of sinusoids of different frequency is injected in the first and the second channel, for the first and the second investigation respectively. These two runs excite the system making use of the two only input channels that we consider in the current model, however the real LTP will count with many other excitation mechanism, like the application of direct forces to the test mass, and hence will require a longer sequence of investigations to fully characterise its dynamics.

3. Model description

The model of the LTP considered here for the analysis is an improved version of the one previously used in previous works [5], which includes actuators and delays. This same model has been considered for similar studies in [6, 7]. In the current analysis, we describe the measurement of an experiment in the LTP in the frequency domain as

$$\vec{\delta} = \mathbf{G} \times \vec{\delta}_{in} + \vec{n},$$

(1)

where $\vec{\delta}$ stands for the measured interferometer displacement; $\vec{\delta}_{in}$ are the injected test signals and $\vec{n}$ the noise vector. All of them two components vectors, one component for each of the measurement channels considered here, the first one measuring the distance between the spacecraft and the first test mass, $o_{x,1}$, and the second one measuring the distance between the two test masses, $o_{x,12}$:

$$\vec{\delta} = \begin{pmatrix} o_{x,1} \\ o_{x,12} \end{pmatrix}, \quad \vec{\delta}_{in} = \begin{pmatrix} o_{in,1} \\ o_{in,12} \end{pmatrix}, \quad \vec{n} = \begin{pmatrix} n_1 \\ n_{12} \end{pmatrix}.$$  

(2)

The transfer functions describing input to output relation in Eq. (1) are contained in the matrix $\mathbf{G}$,

$$\mathbf{G} = (\mathbf{D} \times (\mathbf{DT}_{ifo} \times \mathbf{S})^{-1} + \mathbf{C}_{loop})^{-1} \times \mathbf{C}_{loop} \times \mathbf{DT}_{in}$$

(3)
Figure 1. Box diagram of the model considered in the current analysis. Boxes are systems, rhombus represent injection points and triangles stand for cross-couplings. Notation and symbols are explained in the text.

with

\[ C_{\text{loop}} = C \times A \times H \]  \hspace{1cm} (4)

where we have split the control box, \( C_{\text{loop}} \), in three subsystems: inertia decoupling matrix, \( C \); a matrix with the Drag-Free and Attitude Control Subsystem (DFACS) controllers expressed as transfer functions, \( H \); and the matrix \( A \) with actuator gain and delays. These matrices can be expressed as

\[ C = \begin{pmatrix} - \frac{1}{m_{sc}} & \frac{1}{m_{sc}} \\ 0 & 1 \\ \end{pmatrix}, \]  \hspace{1cm} (5)

\[ H = \begin{pmatrix} H_{df}(\omega) & 0 \\ 0 & H_{sus}(\omega) \end{pmatrix}, \]  \hspace{1cm} (6)

where \( H_{df}(\omega) \) and \( H_{sus}(\omega) \) are (known) transfer function in frequency domain. And,

\[ A = \begin{pmatrix} A_1 e^{-s \tau_1} & 0 \\ 0 & A_2 e^{-s \tau_2} \end{pmatrix}, \]  \hspace{1cm} (7)

where \( A_1 \) and \( A_2 \) are actuator gains for the first —the Field-emission electric propulsion (FEEP)— and the second —the electrostatic capacitive actuators— channel. \( \tau_1 \) and \( \tau_2 \) are delays simulating actuators responses. Default values are \( \tau_1 = 0.2 \) s and \( \tau_2 = 0 \) s.
The dynamics of the test mass are contained in the following matrix

$$
D = \begin{pmatrix}
  s^2 + \omega_1 \left(1 + \frac{m_1}{m_{SC}} + \frac{m_2}{m_{SC}} \frac{\omega_{12}^2}{s^2 + \omega_1 + \omega_{12} - 2 \Gamma_x} \right) \\
  \frac{m_2}{m_{SC}} (\omega_1 + \omega_{12}) + \Gamma_x \\
\end{pmatrix}
$$

(8)

where we define the first and second test mass masses, $m_1$ and $m_2$; the spacecraft mass, $m_{SC}$; first and differential stiffness, $\omega_1$ and $\omega_{12} (= \omega_2^2 - \omega_1^2)$, and gravitational cross-coupling between both test masses, $\Gamma_x$. In this model, the interferometer is represented by a sensing matrix, $S$, translating physical test mass displacements into interferometric read-out,

$$
S = \begin{pmatrix}
  S_{11} & S_{12} \\
  S_{21} & S_{22} \\
\end{pmatrix}
$$

(9)

The four elements are interferometric calibration factors. The non-diagonal are the possible cross-couplings mixing both channels. Processing of interferometer data introduces a delay of 0.4 s on both channels which is indicated by $D_{ifo}$ in the model —see figure [ ] We consider as well a second delay coming from the guidance inputs, this could be useful to compensate for a lack of information on the true time that guidance command is applied inside the DFACS software. Guidance delays are indicated by $D_{in}$ block in the model. Mathematically, both blocks are equivalent:

$$
D_{ifo} = \begin{pmatrix}
  e^{-s \cdot \Delta T_{ifo,1}} & 0 \\
  0 & e^{-s \cdot \Delta T_{ifo,2}} \\
\end{pmatrix}, \quad D_{in} = \begin{pmatrix}
  e^{-s \cdot \Delta T_{in,1}} & 0 \\
  0 & e^{-s \cdot \Delta T_{in,2}} \\
\end{pmatrix}.
$$

(10)

The model is implemented in the LTPDA toolbox as an analytical expression for each of the components. In future version, this model will be implemented as a state-space model in order to enhance its integration with the rest of the models of the toolbox.

4. Description of the parameter estimation methods

In the following we provide a summary of the three methods used in the analysis of the sixth operational exercise. More details for each method can be found in the interleaved references.

**Linear Fit** The procedure for this method is based on a model linearisation to first order. The method performs a time domain parameter estimation. The different steps of the fit process are summarized as follow [ ]:

(i) Data for all investigations and channels are whitened. The whitening filter is obtained through the inverse of the power spectrum for a data segment where no signal is injected, i.e. where we are truly evaluating the noise performance of the instrument.

(ii) Model response to current parameters values is calculated and whitened. This is then subtracted from the data.

(iii) A first order expansion (in terms of the parameters) of the model is calculated and whitened. The set of different series corresponding to different parameters represent the fit basis.

(iv) Each experiment can provide only a limited amount of information, therefore some parameters will be undetermined causing an impair of fit capability. Fit basis is then changed for each experiment by singular value decomposition (SVD) which provides a new fit basis in terms of a linear combination of the elements of the old (physical) basis. The corresponding parameters are linear combination of the original fit parameters.

(v) SVD parameters are estimated for each experiment separately by the solution of the corresponding system of normal equations.
Table 1. Results of the parameter estimation methods for the sixth operational exercise.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Linear</th>
<th>Non-linear</th>
<th>MCMC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\hat{x} \pm \sigma$</td>
<td>$\hat{x} \pm \sigma$</td>
<td>$\hat{x} \pm \sigma$</td>
</tr>
<tr>
<td>A1</td>
<td>1.0699 ± 0.0005</td>
<td>1.0705 ± 0.0006</td>
<td>1.0694 ± 0.0003</td>
</tr>
<tr>
<td>A2</td>
<td>0.99998 ± 0.00003</td>
<td>0.99998 ± 0.00003</td>
<td>0.99996 ± 0.00002</td>
</tr>
<tr>
<td>S21</td>
<td>$(1.2 \pm 0.4) \times 10^{-6}$</td>
<td>$(1.2 \pm 0.4) \times 10^{-6}$</td>
<td>$(1.9 \pm 0.3) \times 10^{-6}$</td>
</tr>
<tr>
<td>$\Delta T_{in,1}$</td>
<td>$-0.1982 \pm 0.0005$</td>
<td>$-0.1985 \pm 0.0005$</td>
<td>$-0.1998 \pm 0.0002$</td>
</tr>
<tr>
<td>$\Delta T_{in,2}$</td>
<td>$-0.199 \pm 0.001$</td>
<td>$-0.199 \pm 0.001$</td>
<td>$-0.199 \pm 0.001$</td>
</tr>
<tr>
<td>$\omega_2^2$</td>
<td>$(-1.319 \pm 0.002) \times 10^{-6}$</td>
<td>$(-1.319 \pm 0.002) \times 10^{-6}$</td>
<td>$(-1.319 \pm 0.002) \times 10^{-6}$</td>
</tr>
<tr>
<td>$\omega_2^2 - \omega_1^2$</td>
<td>$(-7.160 \pm 0.006) \times 10^{-7}$</td>
<td>$(-7.160 \pm 0.006) \times 10^{-7}$</td>
<td>$(-7.150 \pm 0.005) \times 10^{-7}$</td>
</tr>
</tbody>
</table>

(vi) Once the values of all SVD parameters is known with the corresponding uncertainties. The results for the different experiments are collected together and the system is inverted in order to come back to the desired fit parameters.

(vii) New values for fit parameters are then updated and the procedure can start again in a loop until convergence is reached. If the ratio between squared difference of two consecutive parameters estimation and parameters variance is less than 1, then convergence is assumed and the loop is stopped. Parameters which have demonstrated minimum value for mean squared error during the loop are selected for the output.

**Non-linear fit** In this case, the method implements a non-linear parameter estimation in time domain. The key features of the algorithm are [6]:

(i) The joint test statistic, in $\chi^2$ sense, is computed on all available investigations and channels.
(ii) Data are decorrelated by applying whitening filters previously estimated on noisy data stretches.
(iii) To ease the search, parameters are bounded with lower and upper limits.
(iv) The optimization starts from an initial guess by using a derivative-free algorithm (simplex). Eventually, if the minimum is too far away from the optimum, a preliminary large-scale gradient-based algorithm (BFGS Quasi-Newton) is also applied.
(v) Estimated 1-$\sigma$ error are obtained by inverting the Fisher matrix of analytical first-order derivatives.

**Markov Chain Monte Carlo (MCMC)** The MCMC method implemented in LTPDA performs a Monte Carlo integration of the log-likelihood function computed in frequency domain. In summary, the algorithm can be split in following steps [5]:

(i) The Fast Fourier transform is applied to segments where signals are injected and the power spectrum is computed to evaluate the noise level in a signal free segment.
(ii) The expected covariance matrix of the parameters is computed as the inverse of the Fisher matrix.
(iii) The method can optionally look for the maximum likelihood parameters using the simplex algorithm.
(iv) Sample the log-likelihood surface with the Metropolis algorithm. Initially, the method applies an annealing to guarantee a correct exploration of the parameter space. There is a cooling down phase before starting the integration of the likelihood surface around the likelihood
maxima. During this phase, the covariance matrix which defines the proposal distribution is rescaled in some samples to improve convergence. The user sets all parameters that define these phases.

(v) Sample the log-likelihood surface around the maxima obtained in the previous phase until the chain reaches the number of samples set by the user.

5. Parameter estimation methods comparison
The previous methods were applied to the two investigations of the sixth operational exercise. Our Fisher matrix analysis showed that, given the two investigations available, we are only able to estimate 7 parameters independently. We selected the subset shown in table 1, which produces a full rank Fisher matrix. In the table, the results for the three parameter estimation methods show a good agreement for the numerical value as well as for the estimated error on the parameter. Although not shown here, this uncertainty agrees with the one computed from the Fisher matrix analysis —more details on the Fisher matrix computation can be found in [5].

Although the methods seem in good agreement we developed some tools to quantify the agreement and check if the different parameter estimation methods provide us different information about the data being analysed. We discuss this in the next section.

6. Statistical goodness-of-fit tests
We test accuracy of our parameter estimation methods by statistical tests on residuals. Fit residuals are obtained subtracting model response from the measured signal. If the model and the fit process are accurate, then we expect to remove the deterministic signal from our measured signal. Therefore the residuals are expected to be equivalent to the noise output of the system —figures 2 and 3 show the response predicted for each model and the residuals for the two investigations under analysis, respectively.

The main idea behind the test is that subtracting the estimated from the measured response should retrieve a noise curve which is statistically equivalent to the noise level of the system evaluated previously to the signal injection. There are two main limitation to such kind of test: i) The noise out of the system should be stationary and ii) A noise only time series must be available for comparison. The test we construct is based on the comparison of residuals and noise power spectral densities. In order to build a test for power spectral densities we should in general

<table>
<thead>
<tr>
<th>Investigation</th>
<th>Method</th>
<th>Test Result</th>
<th>P-Value Test Statistic</th>
<th>Critical Value @95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inv. 1 oₓ,₁</td>
<td>Linear</td>
<td>Reject</td>
<td>0.0049</td>
<td>0.052</td>
</tr>
<tr>
<td>Inv. 1 oₓ,₁</td>
<td>Non-Linear</td>
<td>Reject</td>
<td>0.0046</td>
<td>0.052</td>
</tr>
<tr>
<td>Inv. 1 oₓ,₁</td>
<td>MCMC</td>
<td>Reject</td>
<td>0.0048</td>
<td>0.052</td>
</tr>
<tr>
<td>Inv. 1 oₓ,₁</td>
<td>Linear</td>
<td>Not Reject</td>
<td>0.6480</td>
<td>0.022</td>
</tr>
<tr>
<td>Inv. 1 oₓ,₁</td>
<td>Non-Linear</td>
<td>Not Reject</td>
<td>0.6480</td>
<td>0.022</td>
</tr>
<tr>
<td>Inv. 1 oₓ,₁</td>
<td>MCMC</td>
<td>Not Reject</td>
<td>0.6543</td>
<td>0.022</td>
</tr>
<tr>
<td>Inv. 1 oₓ,₁</td>
<td>Linear</td>
<td>Not Reject</td>
<td>0.9801</td>
<td>0.014</td>
</tr>
<tr>
<td>Inv. 1 oₓ,₁</td>
<td>Non-Linear</td>
<td>Not Reject</td>
<td>0.9801</td>
<td>0.014</td>
</tr>
<tr>
<td>Inv. 1 oₓ,₁</td>
<td>MCMC</td>
<td>Not Reject</td>
<td>0.9801</td>
<td>0.014</td>
</tr>
<tr>
<td>Inv. 1 oₓ,₁</td>
<td>Linear</td>
<td>Not Reject</td>
<td>0.9801</td>
<td>0.014</td>
</tr>
<tr>
<td>Inv. 1 oₓ,₁</td>
<td>Non-Linear</td>
<td>Not Reject</td>
<td>0.9801</td>
<td>0.014</td>
</tr>
<tr>
<td>Inv. 1 oₓ,₁</td>
<td>MCMC</td>
<td>Not Reject</td>
<td>0.9801</td>
<td>0.014</td>
</tr>
<tr>
<td>Inv. 1 oₓ,₁</td>
<td>Linear</td>
<td>Not Reject</td>
<td>0.9801</td>
<td>0.014</td>
</tr>
<tr>
<td>Inv. 1 oₓ,₁</td>
<td>Non-Linear</td>
<td>Not Reject</td>
<td>0.9801</td>
<td>0.014</td>
</tr>
<tr>
<td>Inv. 1 oₓ,₁</td>
<td>MCMC</td>
<td>Not Reject</td>
<td>0.9801</td>
<td>0.014</td>
</tr>
<tr>
<td>Inv. 1 oₓ,₁</td>
<td>Linear</td>
<td>Not Reject</td>
<td>0.9801</td>
<td>0.014</td>
</tr>
<tr>
<td>Inv. 1 oₓ,₁</td>
<td>Non-Linear</td>
<td>Not Reject</td>
<td>0.9801</td>
<td>0.014</td>
</tr>
<tr>
<td>Inv. 1 oₓ,₁</td>
<td>MCMC</td>
<td>Not Reject</td>
<td>0.9801</td>
<td>0.014</td>
</tr>
<tr>
<td>Inv. 1 oₓ,₁</td>
<td>Linear</td>
<td>Not Reject</td>
<td>0.9801</td>
<td>0.014</td>
</tr>
<tr>
<td>Inv. 1 oₓ,₁</td>
<td>Non-Linear</td>
<td>Not Reject</td>
<td>0.9801</td>
<td>0.014</td>
</tr>
<tr>
<td>Inv. 1 oₓ,₁</td>
<td>MCMC</td>
<td>Not Reject</td>
<td>0.9801</td>
<td>0.014</td>
</tr>
</tbody>
</table>
Figure 2. Investigation 1 in the 6th operational exercise. Top to bottom for channel $o_{x1}$ (left) and $o_{x12}$ (right): Comparison in time domain of the measured output against output predicted by the models resulting from each method (top), residuals in time domain (middle) and residuals in frequency domain compared to independent noise level estimate (bottom).
Figure 3. Investigation 2 in the 6th operational exercise. Top to bottom for channel $o_{x1}$ (left) and $o_{x12}$ (right): Comparison in time domain of the measured output against output predicted by the models resulting from each method (top), residuals in time domain (middle) and residuals in frequency domain compared to independent noise level estimate (bottom).
face the problem of the statistical properties of the spectrum, which, in the case of Welch overlap windowed averaged periodogram, can be highly complex \cite{9} and annoyingly dependent on data underlying distribution. We overcome the problem with the Kolmogorov-Smirnov test (KS-test) whose statistic is not dependent form the statistical properties of the data under test —a more detailed description of this method can be found in \cite{7}.

The null hypothesis for the test is to consider that residuals and noise are two random processes with the same underlying distribution, i.e. residuals and noise are two realizations of the same physical process. The null hypothesis is rejected or not on the basis of a given significance level which is fixed to 5\% in the present case. Null hypothesis is rejected if the probability associated with the KS-statistic of the two data series is less than the required significance level. Results of the KS-test for our methods are reported in table 2. The test provides the same results for the three methods. Spectra of fit residuals for the first channel on the first investigation are not compatible with expected noise spectra, which shows that our methods are not able to completely remove the injected signal in the first channel for this investigation. This effect, that could show a mismatch between the model being used for the analysis and the one used to generate the data, is currently being investigated. In the other 3 experiments the subtraction works effectively and the residuals spectra are compatible with the expected noise spectra.

7. Summary and conclusions

In this contribution we have presented the comparison between three different parameter estimation methods implemented in LTPDA applied to the the analysis of a set of data generated by a complex LTP simulator, the OSE. Results show that the three methods are in agreement when we perform the analysis of the two investigations in the sixth LTP operational exercise.

We also presented a statistical analysis of the residuals obtained by each method when subtracting the expected response of the system to the measured one. All methods agree as well under this framework, however the Kolmogorov-Smirnov analysis show that the methods are not able to completely subtract the injected signal from the measured output in the first channel of the first investigation, i.e. the experiment where we inject a sequence of sinusoids in the first channel. This effect, enhanced by the high signal-to-noise ratio of the signal in this channel, may be due to a mismatch between the model used for the data analysis and the one used for data generation. It is worth reminding here that the OSE is a black-box simulator developed by Astrium GmbH and delivered to ESA under industrial contract.

Next steps of the data analysis team include the improvement of the data analysis methods and focus on the analysis of the different experiments planned during flight operations. In parallel, the models implemented in the LTPDA toolbox to describe the LTP experiment are being implemented as state-space models, which will allow a more modular design and ease the code maintenance and testing.

References

