9^{th} *LISA Symposium, Paris*
ASP Conference Series, Vol. 467
G. Auger, P. Binétruy and E. Plagnol, eds.

Statistical Analysis of LISA Pathfinder Noise

L. Ferraioli

APC, Univ. Paris Diderot, CNRS/IN2P3, CEA/Ifru, Obs de Paris, Sorbonne Paris Cité, France

Abstract. LISA Pathfinder (LPF) is an European Space Agency mission aiming to create the quietest environment for free-falling test masses in order to pave the way to the forthcoming space-based gravitational wave detectors. Reaching such an ambitious target will require a significant amount of system optimisation and characterisation, which will in turn require accurate and quantitative noise analysis procedures. In this paper we present a statistical procedure for the analysis of the noise in spectral domain that is based on the Kolmogorov-Smirnov test. The application on LPF synthetic data reveals the versatility of the Kolmogorov-Smirnov approach, which can easily cope with the correlations introduced by the Welch's overlapped segment averaging method for the calculation of the sample spectrum.

6. Introduction

LISA Pathfinder (LPF) is an European Space Agency mission that will characterize and analyse all possible sources of disturbance which perturb free-falling test masses from their geodesic motion (Antonucci et al. 2011a,b,c; Armano et al. 2009). The system is composed of a single spacecraft (SC) enclosing a scientific payload, the LISA Technology Package (LTP), which is composed of two free-falling test masses (TMs) whose position is sensed by an interferometer. The interferometer senses the relative displacement between the two test masses that represents the main scientific output channel.

One of the final objectives of the LISA Pathfinder mission is the definition of a force noise model for the free-falling test masses. This will require a technique to quantitatively analyse noise data and to assess the differences between noise measurements and models. In this paper we present a technique for the statistical analysis of noise in the spectral domain (Ferraioli et al. 2011). The first difficulty of working in spectral domain is provided by the statistic of the data. In fact spectral data are distributed as a Gamma function. Moreover the method used for the calculation of the sample spectrum is the Welch's overlapped segment averaging (WOSA) method, which introduces two type of correlation in the sample spectrum. The first type of correlation is caused by segment overlapping; it can be attenuated by a proper choice of the overlapping percentage. The second type of correlation is connected with data windowing and it is unavoidable. Our method for noise analysis is capable to overcome all those difficulties in a simple way. It is based on the Kolmogorov-Smirnov (KS) test formalism therefore it works independently of the data underlining statistic. Furthermore, data correlation of the second type can be taken into account in a simple way; it is only influenced by the window applied to the data so a correction factor can be calculated by a Monte Carlo simulation on simple white noise data. The introduction of the correction factor allows to recover the test fairness. All the analysis procedures that we present are available as MATLAB tools in the framework of the LTPDA Toolbox (Hewitson et al. 2009); an object oriented MATLAB Toolbox for advanced data analysis.

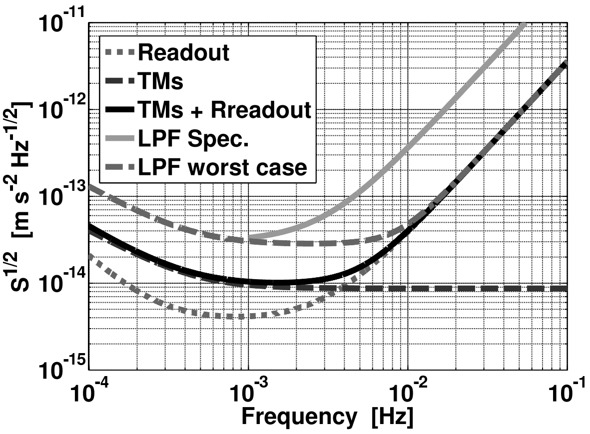

Figure 4. Projection of the spectrum of a_{eff}. 'Readout' is the projection of the readout noise to a_{eff}. 'TMs' is the projection of the force noise on the test masses to a_{eff}. 'TMs + Readout' is the complete noise projection for a_{eff}, it represents the baseline noise level assumed in the present paper. 'LPF worst case' refers to a worst case scenario for a_{eff} and 'LPF Spec.' corresponds to the mission specifications.

7. Application to LISA Pathfinder synthetic data

In this section the procedure for excess noise detection is applied to synthetic LPF data. This provides not only an interesting framework for testing its accuracy and precision, but it also helps clarifying the role of correlations among spectral data.

7.1. Synthetic data and noise projection

LISA Pathfinder is a controlled three body system composed of two test masses and the enclosing spacecraft. One test mass is free falling along the principal measurement axis and it is used as reference for the drag-free controller of the spacecraft. The second test mass is actuated at very low frequencies (below 1 mHz) in order to follow the free falling test mass. This actuation scheme provides a measurement bandwidth $1 \leq f \leq 100$ mHz in which both the test masses can be considered effectively free-falling. The system has two output channels along the principal measurement axis, which are sensing the displacement of the SC relative to the free falling test mass and the relative displacement between the test masses. From the knowledge of the displacement signals and a linear model of the system dynamics, an effective force-per-unit-mass, a_{eff}, acting on the test masses can be extracted. a_{eff} is the combination of the 'true' force-per-unit-mass acting on test masses and a projection of the interferometer readout noise. Following this scheme, in figure 4, the model for the spectrum of the force noise acting on the test masses and the model for the spectrum of the readout noise are both projected into a model for the power spectrum of a_{eff} (TMs + Readout in the figure notation).

In figure 4 we also report the project specifications for LPF and the expected noise spectrum for a_{eff} in a worst case scenario. In our baseline we assumed a reduced force noise on the test masses compared to the worst case but choose to keep the worst case for the readout noise. This was done in order to represent one of the possible scenarios (not the best one) that can be experienced during the mission. The model assumed for the force noise on the test masses is characterized by a low frequency $1/f^2$ behaviour and a flat part for $f > 1$ mHz. The model can be written as $S_{TM}(f) = \theta S^0_{TM}(f)$. θ is an adjustable parameter which takes values $\theta = 1$ for the worst case scenario and $\theta = 0.1$ for our baseline model. It is worth noting that $S_{TM}(f)$

is projected (together with the readout noise model) through the LISA Pathfinder dynamics in order to obtain the expected noise spectrum for a_{eff}, which we indicate with $S_a(f)$. $S_a(f)$ with $\theta = 0.1$ corresponds to 'TMs + Readout' in figure 4, $S_a(f)$ with $\theta = 1$ corresponds, instead, to the 'LPF worst case'.

7.2. Excess noise detection

A change of the noise level on $S_{TM}(f)$ ($\theta \neq 0.1$) produces a variation of the energy content of a_{eff}. In particular we tested the detection of excess noise between two data series and between a data series and a reference model. Synthetic data were produced according to the following procedure:

- Different models for $S_a(f)$ are produced changing θ around the reference value $\theta = 0.1$. Readout noise level is kept fixed.
- Corresponding noise time series for a_{eff} are generated using the procedure reported in (Ferraioli et al. 2010)
- The time series are 24 h long and have a sampling frequency of 1 Hz.
- Sample spectra are calculated for each series with the Welch's overlapped segment averaging algorithm. We chose a Blackman-Harris data window, 50% segment overlap and number of segments averages $N_s = 4$.
- The analysis is restricted to the frequency interval $[0.1, 10]$ mHz since, as can be seen from figure 4, it represents the region in which the force noise on the test masses dominates $S_a(f)$.

The data series for $\theta \neq 0.1$ are compared against the reference series with $\theta = 0.1$. The results of the tests are summarized in table 1. Kolmogorov-Smirnov critical values $d_K(\alpha)$ are calculated for a significance level $\alpha = 0.05$ which corresponds to a 95% confidence. Each value of θ corresponds to a value of the in-band energy content $E(\theta)$ of $S_a(f)$ in the analysed frequency band. We report in table 1 the relative change in energy $\Delta E/E$ corresponding to a relative change in θ. Spectral data are correlated among different frequency values because of (Percival & Walden 1993; Thomson 1977):

- the windowing of the time series that corresponds to a convolution in the frequency domain of the window function with the sample spectrum and
- the overlap between data segments in the Welch's overlapped segment averaging algorithm.

The first effect is unavoidable, the second effect, instead, can be attenuated by a proper choice of the segment overlap percentage (50% overlap for a Blackmann-Harris window).

In the Kolmogorov-Smirnov test an empirical cumulative distribution (ECDF) is tested against a continuous theoretical model or, alternatively, two ECDFs are tested with the hypothesis that they share the same limiting cumulative distribution function. The standard critical values, $d_K(\alpha)$, are valid only if the correlations among data are negligible. If this is not the case, the effective statistic of the estimator can be numerically calculated with Monte Carlo simulations. The corresponding results of a Monte Carlo simulation with $N_{MC} = 5000$ realizations of the reference data series are indicated in table 1 with the suffix MC. In the same table, the symbol $\surd$ is used to indicate that the spectral data for the corresponding $\Delta E/E$ is compatible with the reference. On the contrary the symbol $\times$ indicates a rejection.

Observing the test results reported in columns 7 and 9, the comparison of the data with $d_K(\alpha)$ determines a rejection for $\Delta E/E = 0.05$. On the other hand, the same value is accepted when Monte Carlo result $d_K^{MC}(\alpha)$ is used. In this case the presence of correlation among data has affected the tests statistic and therefore the standard equation, that assumes independence among data, is not valid.

Looking at the results for the Kolmogorov-Smirnov test between two data series, we discover that $d_K(\alpha)$ and $d_K^{MC}(\alpha)$ are practically equal. In fact the results for the two corresponding

Table 1. Detection of noise differences in the frequency band [0.1, 10] mHz. The symbol √ indicates compatibility between tested objects. The symbol × is instead used for indicating test rejection. θ is an adjustable parameter which assumes value $\theta = 0.1$ for our baseline model. Different values of θ correspond to different values of the in-band energy content $E(\theta)$ of $S_a(f)$. We reported here relative values with respect to the baseline reference.

		KS vs. data			KS vs. model			
			$d_K(\alpha)$	$d_K^{MC}(\alpha)$		$d_K(\alpha)$	$d_K^{eff}(\alpha)$	$d_K^{MC}(\alpha)$
$\frac{\Delta \mathbf{E}}{\mathbf{E}}$	$\frac{\Delta\theta}{\theta}$	d_K	0.1030	0.1006	d_K	0.0730	0.0982	0.0969
−0.14	−0.7	0.2352	×	×	0.2547	×	×	×
−0.08	−0.4	0.1424	×	×	0.1103	×	×	×
−0.06	−0.3	0.1806	×	×	0.1747	×	×	×
−0.05	−0.25	0.1740	×	×	0.1685	×	×	×
−0.04	−0.2	0.0887	√	√	0.0696	√	√	√
−0.03	−0.15	0.0906	√	√	0.0433	√	√	√
−0.01	−0.05	0.0771	√	√	0.0302	√	√	√
0.01	0.05	0.0551	√	√	0.0539	√	√	√
0.03	0.15	0.0629	√	√	0.0370	√	√	√
0.04	0.2	0.0603	√	√	0.0635	√	√	√
0.05	0.25	0.0519	√	√	0.0818	×	√	√
0.06	0.3	0.0857	√	√	0.1253	×	×	×
0.08	0.4	0.1103	×	×	0.1137	×	×	×
0.14	0.7	0.1520	×	×	0.2423	×	×	×

columns of table 1 (columns 4 and 5) are in perfect agreement. This is the practical result of one of the most interesting properties of the Kolmogorov-Smirnov test. Since the sample spectra in our test are calculated following the same procedure, the underlying distribution is the same so the test statistic is not spoiled.

Comparing $d_K(\alpha)$ and $d_K^{MC}(\alpha)$ of columns 7 and 9 we see that the effect of correlation is to increase the maximum expected spread between the ECDF and limiting cumulative function. Therefore the effect of data correlation is to distort the expected statistic for d_K. The ‘distortion’ of d_K statistic can be taken into account if an effective value for the parameter K is introduced. In the case of the comparison of the ECDF for correlated data against a theoretical cumulative function the application of the standard values for d_K ($K = N_f$, N_f being the number of data elements) leads to a statistically unfair test. We then discovered that the test fairness can be recovered if an effective value for K is used rather than the standard $K = N_f$. In particular, for spectral data produced with the Welch’s overlapped segment averaging method, Blackmann-Harris window, $N_s = 4$ averages on 50% overlapped segments, we obtained $K_{eff} = \beta N_f$ with $\beta = 0.55$ for a significance level $\alpha = 0.05$. It is worth noting that the value of β is independent from the number of data points considered, and from the spectral shape, provided that the different shapes have reasonably comparable smoothness on a frequency interval comparable with the width of the first lobe of the data window. As an example, the value of $\beta = 0.55$ is valid for LPF-like data and for white-noise equivalently.

8. Conclusions

The problem of excess noise detection and noise parameter estimation for non-Gaussian data is analysed in the framework of the LISA Pathfinder mission. Excess noise detection can be approached in two ways. In one way, the noise content of a data series is compared with a reference data series, in the other way the noise content of the data is compared with a reference model. The Kolmogorov-Smirnov estimator, that we propose, has the advantage of being independent of the statistical properties of the data under test, this is an advantageous properties since the sample spectrum is distributed as a Gamma function. Furthermore, we demonstrated that the introduction of a shape parameter allows to use standard equations to calculate proper boundaries for Kolmogorov-Smirnov confidence intervals. The shape parameter depends on the required significance level and on the data correlations. It does not depend on the number of data considered for the test. Since correlation among spectral data is mainly introduced by the windowing process, the shape parameter is fixed for a wide family of spectral shapes. For example, synthetic LISA Pathfinder data share the same shape parameter with white noise.

Acknowledgments. The author wishes to acknowledge the CNES for its financial support.

References

Antonucci, F., Armano, M., Audley, H., Auger, G., Benedetti, M., Binetruy, P., Boatella, C., Bogenstahl, J., Bortoluzzi, D., Bosetti, P., Brandt, N., Caleno, M., Cavalleri, A., Cesa, M., Chmeissani, M., Ciani, G., Conchillo, A., Congedo, G., Cristofolini, I., Cruise, M., Danzmann, K., Marchi, F. D., Diaz-Aguilo, M., Diepholz, I., Dixon, G., Dolesi, R., Dunbar, N., Fauste, J., Ferraioli, L., Fertin, D., Fichter, W., Fitzsimons, E., Freschi, M., a Marin, A. G., a Marirrodriga, C. G., Gerndt, R., Gesa, L., Giardini, D., Gibert, F., Grimani, C., Grynagier, A., Guillaume, B., GuzmÃ!'n, F., Harrison, I., Heinzel, G., Hewitson, M., Hollington, D., Hough, J., Hoyland, D., Hueller, M., Huesler, J., Jeannin, O., Jennrich, O., Jetzer, P., Johlander, B., Killow, C., Llamas, X., Lloro, I., Lobo, A., Maarschalkerweerd, R., Madden, S., Mance, D., Mateos, I., McNamara, P. W., Mendes, J., Mitchell, E., Monsky, A., Nicolini, D., Nicolodi, D., Nofrarias, M., Pedersen, F., Perreur-Lloyd, M., Perreca, A., Plagnol, E., Prat, P., Racca, G. D., Rais, B., Ramos-Castro, J., Reiche, J., Perez, J. A. R., Robertson, D., Rozemeijer, H., Sanjuan, J., Schleicher, A., Schulte, M., Shaul, D., Stagnaro, L., Strandmoe, S., Steier, F., Sumner, T. J., Taylor, A., Texier, D., Trenkel, C., Tombolato, D., Vitale, S., Wanner, G., Ward, H., Waschke, S., Wass, P., Weber, W. J., & Zweifel, P. 2011a, Classical and Quantum Gravity, 28, 094002. URL `http://stacks.iop.org/0264-9381/28/i=9/a=094002`

Antonucci, F., Armano, M., Audley, H., Auger, G., Benedetti, M., Binetruy, P., Boatella, C., Bogenstahl, J., Bortoluzzi, D., Bosetti, P., Caleno, M., Cavalleri, A., Cesa, M., Chmeissani, M., Ciani, G., Conchillo, A., Congedo, G., Cristofolini, I., Cruise, M., Danzmann, K., Marchi, F. D., Diaz-Aguilo, M., Diepholz, I., Dixon, G., Dolesi, R., Dunbar, N., Fauste, J., Ferraioli, L., Fertin, D., Fichter, W., Fitzsimons, E., Freschi, M., a Marin, A. G., a Marirrodriga, C. G., Gerndt, R., Gesa, L., Gilbert, F., Giardini, D., Grimani, C., Grynagier, A., Guillaume, B., GuzmÃ!'n, F., Harrison, I., Heinzel, G., Hewitson, M., Hollington, D., Hough, J., Hoyland, D., Hueller, M., Huesler, J., Jeannin, O., Jennrich, O., Jetzer, P., Johlander, B., Killow, C., Llamas, X., Lloro, I., Lobo, A., Maarschalkerweerd, R., Madden, S., Mance, D., Mateos, I., McNamara, P. W., Mendes, J., Mitchell, E., Monsky, A., Nicolini, D., Nicolodi, D., Nofrarias, M., Pedersen, F., Perreur-Lloyd, M., Perreca, A., Plagnol, E., Prat, P., Racca, G. D., Rais, B., Ramos-Castro, J., Reiche, J., Perez, J. A. R., Robertson, D., Rozemeijer, H., Sanjuan, J., Schleicher, A., Schulte, M., Shaul, D., Stagnaro, L., Strandmoe, S., Steier, F., Sumner, T. J., Taylor, A., Texier, D., Trenkel, C., Tombolato, D., Vitale, S., Wanner, G., Ward, H., Waschke, S., Wass, P., Weber, W. J., & Zweifel, P. 2011b, Classical and Quantum Gravity, 28, 094001. URL `http://stacks.iop.org/0264-9381/28/i=9/a=094001`

Antonucci, F., Armano, M., Audley, H., Auger, G., Benedetti, M., Binetruy, P., Boatella, C., Bogenstahl, J., Bortoluzzi, D., Bosetti, P., Caleno, M., Cavalleri, A., Cesa, M., Chmeissani, M., Ciani, G., Conchillo, A., Congedo, G., Cristofolini, I., Cruise, M., Danzmann, K., Marchi, F. D., Diaz-Aguilo, M., Diepholz, I., Dixon, G., Dolesi, R., Fauste, J., Ferraioli, L., Fertin, D., Fichter, W., Fitzsimons, E., Freschi, M., a Marin, A. G., a Marirrodriga, C. G., Gesa, L., Giardini, D., Grimani, C., Grynagier, A., Guillaume, B., GuzmÃ!'n, F., Harrison, I., Heinzel, G., Hewitson, M., Hollington, D., Hough, J., Hoyland, D., Hueller, M., Huesler, J., Jeannin, O., Jennrich, O., Jetzer, P., Johlander, B., Killow, C., Llamas, X., Lloro, I., Lobo, A., Maarschalkerweerd, R., Madden, S., Mance, D., Mateos, I., McNamara, P. W., Mendes, J., Mitchell, E., Monsky, A., Nicolini, D., Nicolodi, D., Nofrarias, M., Pedersen, F., Perreur-Lloyd, M., Perreca, A., Plagnol, E., Prat, P., Racca, G. D., Rais, B., Ramos-Castro, J., Reiche, J., Perez, J. A. R., Robertson, D., Rozemeijer, H., Sanjuan, J., Schulte, M., Shaul, D., Stagnaro, L., Strandmoe, S., Steier, F., Sumner, T. J., Taylor, A., Texier, D., Trenkel, C., Tombolato, D., Vitale, S., Wanner, G., Ward, H., Waschke, S., Wass, P., Weber, W. J., & Zweifel, P. 2011c, Classical and Quantum Gravity, 28, 094006. URL http://stacks.iop.org/0264-9381/28/i=9/a=094006

Armano et al., M. 2009, Class. Quant. Grav., 26, 094003

Ferraioli, L., Congedo, G., Hueller, M., Vitale, S., Hewitson, M., Nofrarias, M., & Armano, M. 2011, Phys. Rev. D, 84, 122003. URL http://link.aps.org/doi/10.1103/PhysRevD.84.122003

Ferraioli, L., Hueller, M., Vitale, S., Heinzel, G., Hewitson, M., Monsky, A., & Nofrarias, M. 2010, Phys. Rev. D, 82, 042001

Hewitson, M., Armano, M., Benedetti, M., Bogenstahl, J., Bortoluzzi, D., Bosetti, P., Brandt, N., Cavalleri, A., Ciani, G., Cristofolini, I., Cruise, M., Danzmann, K., Diepholz, I., Dolesi, R., Fauste, J., Ferraioli, L., Fertin, D., Fichter, W., Garcia, A., Garcia, C., Grynagier, A., Guzman, F., Fitzsimons, E., Heinzel, G., Hollington, D., Hough, J., Hueller, M., Hoyland, D., Jennrich, O., Johlander, B., Killow, C., Lobo, A., Mance, D., Mateos, I., McNamara, P. W., Monsky, A., Nicolini, D., Nicolodi, D., Nofrarias, M., Perreur-Lloyd, M., Plagnol, E., Racca, G. D., Ramos-Castro, J., Robertson, D., Sanjuan, J., Schulte, M. O., Shaul, D. N. A., Smit, M., Stagnaro, L., Steier, F., Sumner, T. J., Tateo, N., Tombolato, D., Vischer, G., Vitale, S., Wanner, G., Ward, H., Waschke, S., Wand, V., Wass, P., Weber, W. J., Ziegler, T., & Zweifel, P. 2009, Classical and Quantum Gravity, 26, 094003. URL http://stacks.iop.org/0264-9381/26/i=9/a=094003

Percival, D. B., & Walden, A. T. 1993, Spectral Analysis for Physical Applications (Cambridge, UK: Cambridge University Press)

Thomson, D. J. 1977, Bell Sys. Tech. Journal, 56, 1769