University of Trento

Department of Physics

Residual test mass acceleration in LISA Pathfinder: in-depth statistical analysis and physical sources

Author:
Lorenzo Sala

Supervisor:
Prof. Stefano Vitale

A thesis submitted for the degree of
Doctor of Philosophy in Physics

Doctoral cycle XXXV

July 17, 2023

"You see, even wrong ideas can be nice"
— José Saramago, *The Cave*

Abstract

LISA Pathfinder (LPF) has been a space mission led by ESA with NASA contributions, operating between March 2016 and July 2017. LPF demonstrated the feasibility of setting bodies in space along freely-falling geodetic trajectories, complying with the residual acceleration requirements of the future gravitational-wave observatory LISA.
LISA Pathfinder represented an important milestone in testing the technological readiness of many instruments which will fly onboard LISA. In particular, it tracked with sub-pm interferometric precision the distance between two test masses, measuring the stray relative acceleration between them. Having low stray acceleration noise at low frequencies (less than $20\,\mathrm{fN\,Hz^{-1/2}}$ at 0.1 mHz) is paramount for LISA, as it directly impacts gravitational wave detection and observation. After operations, the LPF Collaboration pointed out that two phenomena, affecting the sub-mHz performance, were not completely understood and needed deeper analyses. This, despite performing better than requirements. Such phenomena are, namely, the low-frequency acceleration noise, and the sub-pN transient acceleration glitches. This thesis work focuses entirely on analyzing these observations, in view of the future mission LISA.

Regarding the low-frequency sub-mHz noise, first, we make a preliminary analysis. We investigate its evolution in time, its properties, its stability, and its nature. Informed by results from previous analyses, we define useful figures of merit to analyze the noise stability over the LPF mission. We find that the low-frequency noise has had a remarkably stable behavior for nearly two years, but noise fluctuations are not compatible with an overall unique noise.
We develop decorrelation tools to understand the measured noise's physical origin. In particular, LPF measured several environmental quantities, – magnetic fields, temperature, thermal gradients, etc. – synchronously to the test mass acceleration. We aim at finding, if any, correlations between those time series and the main acceleration measurements. Then we give an overview of other known and quantifiable effects, summarizing their impact on acceleration.

We expand previous analyses about the LPF outgassing environment. Pressure, on LPF, is indirectly measured as stochastic acceleration noise generated by random Brownian collisions between the residual gas and the test masses. The measurement is so precise that the measured white-noise level is effectively a good proxy for pressure. We analyze pressure evolution, fitting data to models. We also analyze the long-term quasi-static acceleration drift observed on LPF, proposing a physical model.

A large part of this thesis is dedicated to statistical work on the

properties of stochastic processes, and applications to the LPF case. We develop results on multivariate spectral estimation. Implementing results from complex-variable statistics, we show that cross-power spectral density matrices follow complex-Wishart probability distributions, and develop a Bayesian tool for the posterior inference of spectral parameters.

We extensively analyze the second phenomenon impacting low-frequency performances, the acceleration transient glitches. First, we characterize their physical properties. We show that LPF glitches spanned a wide range of amplitudes, transferring impulses between a few fN s, to some nN s, and showing durations ranging from a few seconds to hours.
We show that LPF glitches fall into two rather distinct categories: fast transients in the interferometric motion readout and long-lasting sub-pN true force transient events, acting on the test masses. We present an analysis of the physical and statistical properties of both, including a cross-investigation with other time series such as magnetic fields, temperature, differential torque, and other dynamical variables, and examine the possible sources of the force glitches, identifying the most likely ones.

Contents

List of Figures

List of Tables

List of Acronyms

1TMTP Single-test-mass torsion pendulum.

4TMTP Four-test-mass torsion pendulum.

CMS Charge Management System.

CPSD Cross Power Spectral Density.

CR Cold Runs (See Table A.1 on page 193).

CVM Caging and Venting Mechanism.

DAL Days after launch.

DAV Days After Venting.

DFT Discrete Fourier Transform.

DOY Day of year.

DRS Disturbance Reduction System.

FLG Fast, low-impulse glitch.

GRS Gravitational Reference Sensor.

GW Gravitational Wave.

IBM Internal Balance Mass.

ICG Impulse-carrying glitch.

IFO Interferometer.

LEO Low Earth orbit.

LISA Laser Interferometer Space Antenna.

LPF LISA Pathfinder.

LTP LISA Technology Package.

LTPDA LTP Data Analysis toolbox.

LVK LIGO-Virgo-KAGRA.

MCMC Markov Chain Monte Carlo.

MOSA Moving Optical Sub-Assembly.

MSC Magnitude Square Coherence.

OB Optical Bench.

OMS Optical Metrology System.

OR Ordinary Runs (See Table A.1 on page 193).

PSD Power Spectral Density.

PTA Pulsar Timing Array.

SC Spacecraft.

SNR Signal-to-Noise Ratio.

TDI Time Delay Interferometry.

TP Torsion pendulum.

VE Vacuum Enclosure.

Preface

This thesis work is meant to give an in-depth analysis of the acceleration performance of the LISA Pathfinder (LPF) mission. LPF, led by the European Space Agency with NASA contributions, operated between 2015 and 2017, succeeding in its objectives and proving that the relative acceleration between free-falling bodies can be kept low enough to meet the requirements for the observation of gravitational waves.
The LPF mission was intended as a precursor of the LISA mission, the Laser Interferometer Space Antenna, ESA's gravitational waves observatory, which will fly in the mid-2030s. Its main purpose was to set two test masses into free fall, to within a few $\mathrm{fm\,s^{-2}\,Hz^{-1/2}}$ spurious acceleration at mHz frequencies.

The first paper about the noise performance [1] was published in 2016, already showing that LPF met not only its requirements, but also the more stringent requirements for LISA. Analyses in that paper referred to the noise run Apr16A of Table A.1, which will be explained later in the thesis but, as can be noticed, was one of the first runs performed in the mission.
Many runs, experiments, and analyses were performed ever since, resulting in another publication about noise performance [2], which was published after the mission ending, and already contained details about all the noise runs. That reference focuses on the "best" noise run Feb17B, which is usually taken as a reference run for LPF, for reasons that will also be clearer in this work. In that publication, it was already noted that the LPF low-frequency performance was affected by two issues, possibly different in nature: a low-frequency noise spectrum showing an excess above predictions, and the occurrence of transients, carrying impulses from fm/s to nm/s, showing up with a rate 1/day during ordinary runs.

Work and analyses about the LPF noise performance went on. In her PhD work, Eleonora Castelli [3] carried out analyses on a more complete set of LPF runs, summarizing the post-processing performed to produce data, and providing analyses of the results.

The work here presented is built on the three cited works, extending the analyses in [3] with in-depth statistical investigations. Starting from the post-processed experimental runs, we analyze the detected acceleration, its time-domain and frequency-domain features, and develop statistical tools meant to better explore the physical origin of the observed noise.
Within this work, we published the paper in reference [4], regarding LPF transients. Moreover, another work regarding low-frequency noise, its properties, its stability, its

evolution, and its physical origin, is currently in preparation [5].

In Chapter 1 we give a qualitative description of the LISA mission, the gravitational-wave detection principle, and the onboard instrumentation. This introduction is meant to set the context in which LISA Pathfinder was developed, to better understand the scientific objectives of LPF.

In Chapter 2 we give more details about the LPF instrument and operations. In the first section, we give a broad description of the mission, the onboard instrumentation, and the main issues concerning acceleration noise performance. In the second section, we describe the on-ground data processing, the building process of the dynamical variables, and the subdivision of the LPF runs.

Chapter 3 is a bit off-track from LPF analyses. In that chapter we present the spectral estimation tools that we used, starting from the basics. All the work presented there is essential for the following chapters. In addition to the commonly-used tools, we present multivariate time series spectral analysis with a statistical approach, emphasizing the probability distributions of spectral quantities. We describe novel Bayesian methods for time series decorrelation, which we will employ in later chapters and in the analyses in [5].

In Chapter 4 we start by focusing on the sub-mHz noise branch, distinguishing the low-frequency branch from the white-noise plateau, which we analyze separately. Analyzing a selected set of noise-only runs, we focus on the low-frequency branch, analyzing the noise stability, its stationarity, its evolution in time, and its nature. Work in this chapter is used for analyses in [5].

In Chapter 5 we instead focus on the white-noise plateau, showing that it is dominated by Brownian noise. Then, we analyze its evolution, inferring parameters of the outgassing environment of LPF's vacuum chambers. Work in this chapter will be published in [5].

In Chapter 6 we go on with the analysis of the low-frequency noise, cross-correlating it with many LPF telemetries, to understand its physical nature. We employ the algorithms developed in Chapter 3 to decorrelate time series, and finally give a breakdown projection of all the known contributions to force noise. Work in this chapter will be published in [5].

Chapter 7 is fully dedicated to the LPF transient events, known as glitches, which occurred all over the LPF mission. Closely following the work we published in 2022 [4], we first present the classification of glitches, the observed parameters, and the observed glitch nature. Then, after cross-correlating with other dynamical variables and environmental variables, we present deeper analyses on their nature.

In Chapter 8 we give a final examination of the results found so far, with discussions on the possible physical origin of the detected excess noise and transient glitches.

1 The LISA Mission

The search for Gravitational Waves from space

LISA, the Laser Interferometer Space Antenna, will be the first space-based observatory to detect gravitational waves (GW). In 2017, ESA selected LISA as one of the three large-class missions (L3) within the *Cosmic Vision* program, planning it for launch in the mid-2030s [6, 7]. LISA will enable the observation of low-frequency GWs, expanding the observational window of ground-based detectors toward lower frequencies. Currently, the ground-based detectors from the LVK Collaboration (LIGO-Virgo-KAGRA) allow the measurement of gravitational waves in the frequency band 10 Hz–10 kHz. LISA will have a different spectrum, providing a goal observational band[1] 20 µHz–1 Hz, with a sweet spot at 10 mHz.

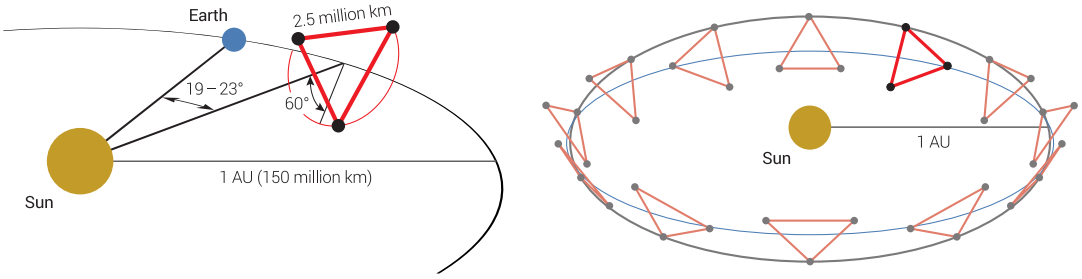


Figure 1.1: Concept representation of the orbits of the LISA constellation. The three spacecraft will follow an Earth-trailing orbit, following the Earth around the Sun at about 20° distance. Adapted from [8].

From the astrophysical point of view, having access to a lower frequency band has a significant impact. Not only does it allow to improve observation capabilities, but it opens to a completely different astrophysical scenario. LISA will observe GWs from many different astrophysical bodies, from merging massive black holes, to inspiralling celestial bodies, both within the Milky Way and outside of it.
Low-frequency GW observation will also change the concept of GW observation itself. Because of GW emission laws, the lower the orbital frequency is, the longer the source emits GWs at that frequency. For this reason, the typical LISA source produces a peaked spectral line, increasing its frequency slowly during the inspiral. Some sources

[1] There is a substantial difference between requirements and goals. Driven by science objectives and technical feasibility, requirements represent the experimental needs, necessary to achieve the objectives. Goals, more relaxed than requirements, are instead technically more difficult to achieve, and not strictly needed to be met. However, science would greatly benefit if requirements were overcome, possibly reaching goals. For LISA, the frequency band requirement is 100 µHz–1 Hz, whereas the goal band is wider, and it is currently set at 20 µHz–1 Hz.

could even inspiral emitting in the LISA band for months or years. LISA will therefore observe many emitting sources at the same time, superimposing and overlapping, without even the possibility of measuring the instrument noise alone.
This short chapter is dedicated to a broad introduction to the objectives of LISA, the mission concept, and the key instruments. This introduction is not meant to be complete about the LISA mission; the purpose is to set the context of the LISA Pathfinder mission, its experimental relevance, and its meaning, as that is the main focus of this work. Many of the common concepts between LISA and LISA Pathfinder are just introduced here, and better explained in the next chapter, which is dedicated to LISA Pathfinder.

1.1 The mission concept

The LISA mission will be a constellation of three identical spacecraft. These spacecraft will orbit the Sun at 1 astronomical unit distance, trailing the Earth along its orbit, about 20° behind it. The nominal inter-spacecraft distance will be 2.5 million km, enabling the long-arm measurements for GW detection. As shown in Fig. 1.1 on the previous page, the spacecraft will maintain a nominal equilateral triangle configuration, whose orientation in space will change over time. This variation is at the basis of the mission principle, at least for two reasons. First, orbital mechanics requires no active thrust to maintain the triangular-shaped configuration: this allows to effectively have reliable geodetic references, as better explained later. Secondly, changing the orientation also changes the antenna sensitivity pattern, allowing to precisely pinpoint the location of GW sources during their observation, and to improve the sensitivity to different GW polarizations.

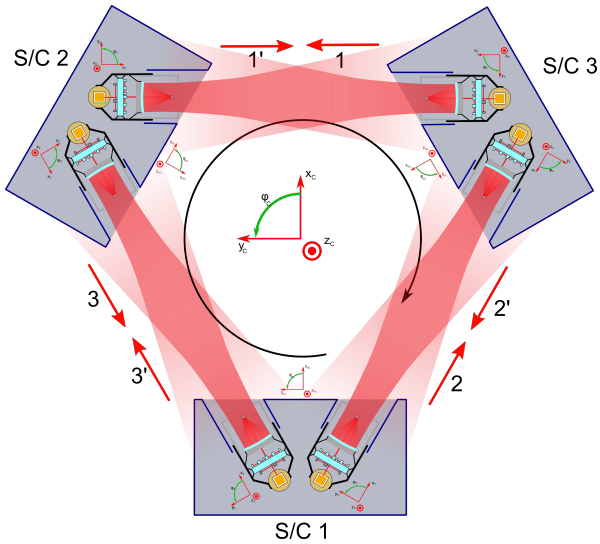


Figure 1.2: LISA constellation concept. The test masses are schematically represented within each spacecraft, virtually connected by laser beams for interferometry. Adapted from [8].

The key elements, within each spacecraft, will be the geodesic reference Au-Pt test masses (TM), as represented in Fig. 1.2. Each spacecraft will be equipped with two of

these masses, which will act as free-falling references. Each TM will face an opposite spacecraft, ideally behaving as the end mirror of a very-long-arm interferometer.
Spacecraft dynamics is a complex subject, as its task is to ensure that the TMs are moving along geodesics along the inter-spacecraft axes. Along the sensitive interferometric axes, the TMs will be free to move in space, and the spacecraft will follow their motion with a *drag free* configuration: micro-thrusts will be applied to the spacecraft themselves, to adjust their trajectories and maintain their nominal positions with respect to the TMs. Along the non-sensitive degrees of freedom, it will be the TMs to follow the spacecraft, thanks to the application of electrostatic actuation forces.
The TMs will however have no mechanical contact with the surrounding electrodes and structure: the GW signals to be detected are so faint that the noise generated by a mechanical contact would hinder the detection.

1.2 The detection principle

The detection principle relies on the laws of general relativity. As reported in every standard introductory textbook about general relativity, e.g. [9], gravitational waves, after being emitted for instance by astrophysical events, propagate in spacetime at the speed of light.

1.2.1 Gravitational-wave physics

The observable effect of passing gravitational waves is a very weak deformation of spacetime itself, which ultimately is an alteration of the proper space-time distance between free-falling observers. In this sense, the test masses act in the LISA case as the free-falling references, following geodetic trajectories, ideally undisturbed in absence of GWs.
The spacetime deformation is known under the name of *gravitational strain*, $h(t)$, which acts with two polarizations. The strain acts as a dilation coefficient, inducing a variation of the proper distance L between two observers by an amount ΔL, which is proportional to L itself. In the case of optimal incidence direction and polarization, $h \sim \Delta L/L$. To give a scale of the phenomenon, the (adimensional) strain from distant astrophysical events is of the order of 10^{-20}. Such a small relative length variation requires on the one hand long-arm measurements, as long as possible; on the other hand, it requires high-precision end-to-end measurements, which nowadays are only achievable with interferometers. In the 3-4 km interferometers of the LVK Collaboration, a partial improvement of the overall sensitivity is achieved with high-power lasers, reflecting in Fabry-Pérot optical cavities [10].
Interferometry on LISA will be inherently different from the ground-based one: first of all, because of the long inter-spacecraft distance; second, because spacecraft move with high relative velocity. Later, we will give more details about how interferometry will be performed.

1.2.2 The need for a space mission

In describing the LISA mission, we still have not touched the most important point, that is, why a space mission is needed and why the same results can not be achieved with on-ground detectors. To discuss this, we first have to tell a bit more about how GW interferometric detection works, with a simple model of a LISA laser link.

The key point is that during the propagation of the laser beam from the emitting spacecraft to the receiving one, any passing GWs act all over time, imprinting their waveforms as phase shifts of the laser beam. When received by the distant spacecraft, the beam phase is compared to a local laser beam and measured. In the case of optimal incidence and polarization, the phase shift reads

$$\Delta\phi(t_{\mathrm{rec}}) = \frac{2\pi}{\lambda_{\mathrm{Laser}}}\frac{c}{2}\int_{t_{\mathrm{em}}}^{t_{\mathrm{rec}}} h(t)\,dt \tag{1.1}$$

If the propagation time L/c is smaller than the characteristic time of the GW strain, then the expression for the phase reduces to

$$\Delta\phi(t_{\mathrm{rec}}) \sim \frac{2\pi}{\lambda_{\mathrm{Laser}}}\frac{L}{2}\,h(t_{\mathrm{rec}}) \tag{1.2}$$

This view of the effects of GW on laser light is not the only one possible. Equivalently, one could indeed change their point of view, and see the effect of GWs on the propagation of laser beams as a variation of frequency, $\Delta\nu(t) \propto \dot{h}(t)$. From this perspective, the effect of GWs is equivalent to a Doppler shift. It is then indistinguishable from a real Doppler shift, caused for example by a spurious acceleration of the interferometer end-points.
In this sense, real forces disturbing the interferometer end-points are in direct competition with GWs, since they generate the same observable effect. Force noise affecting the end mirrors is listed as a primary noise source.

In ground-based interferometers, the motion of the end mirrors is inherently affected by the seismic and Newtonian noise [10]. Respectively, seismic and Newtonian noises are caused by Earth's micro-seismic activity, and local fluctuations of the gravitational field. These noises are the origin of a steep feature in the spectral sensitivity of such detectors, which prevents measurements below $\sim 10\,\mathrm{Hz}$ frequencies. The Earth's environment prevents GW detection at low frequencies, requiring to move to space.

1.2.3 The need for a technology-readiness test – LPF

In space, the situation is much better because of the absence of seismic and Newtonian noise. Nevertheless, other issues arise. For example, other acceleration sources become limiting. In modeling LISA at first, these sources were in principle possibly unknown and unmodeled, and for the same on-ground limitations as before could not be fully measured and tested on Earth.
One major noisy force in space is known to be the radiation pressure from the Sun, which fluctuates at low frequencies, especially those of interest for GW observation. Other external sources include interaction with the interplanetary magnetic fields,

and interaction with cosmic rays. Nevertheless, most of the force sources could also come from the spacecraft itself.

Mainly for these reasons, ESA decided to launch another mission, LISA Pathfinder (LPF) to test the technology readiness for LISA. Instead of having the end-point test masses in two different spacecraft, LPF was a single spacecraft with two test masses, whose distance was tracked by a sub-pm interferometer. The 2.5 million km arm length was shrunk to 38 cm, preventing any measurement of GWs, but allowing the measurement of stray acceleration noise.

LPF succeeded in its objectives, proving that stray acceleration noises can be kept low enough to allow the observation of GWs, complying with the requirements for LISA [1–3]. Acceleration noise during the best-performance run is shown in Fig. 2.17 on page 35, in the lower panel, along with the LISA requirements. LPF also succeeded in testing many technologies for LISA. More details and references will be given in the next chapter, but, making a non-exhaustive list, LPF:

- Tested the first sub-pm interferometer ever operated in space, and the operability in space of its optical bench.
- Tested the contact-free UV-light Charge Management System, which will also operate on LISA with some adaptations.
- Provided a full test of the LISA Gravitational Reference System, the actuation controls, and the sub-femto-g drag-free control loops.
- Provided a test for the force-noise induced by electrostatic actuation and sensing.
- Tested the in-flight TM caging, release, and grabbing mechanisms, providing launch-lock mechanisms and geodesic-insertion once on orbit.
- Tested the magnetic, thermal, and pressure environment, and analyzed their impact on acceleration noise performance.

1.3 LISA instrument overview and requirements

Broadly speaking, LISA will provide high-precision interferometry between three pairs of test masses, floating freely within different spacecraft. Each pair of masses will constitute a two-way laser link, which ultimately provides the key measurements of the mission, as the information on the gravitational strain will be embedded in the laser phase shift.
Single-link measurement will not be performed with a single laser beam, but the measurement will in practice be split into three parts: a local TM-to-SC measurement, a long-arm high-power SC-to-SC measurement, and another local SC-to-TM measurement, as in Fig. 1.3. Backlink fibers will allow laser beams to be shared by different laser links, necessary to perform interferometry along different arms.

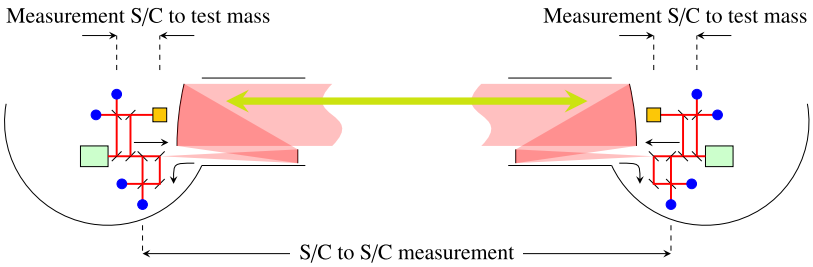


Figure 1.3: Partition of the interferometric LISA measurement, simplified scheme. The total TM-to-TM measurement is virtually split into two local TM-to-SC measurements and a long-arm SC-to-SC measurement. Adapted from [11].

The spacecraft will follow a quasi-equilateral triangular configuration following Earth's orbit, keeping a constant distance from the Earth and the Sun. The orbits will not need active thrust to be maintained, keeping the quasi-equilateral configuration as a result of orbital dynamics. Nevertheless, the angle between the spacecraft, which is nominal 60°, is not perfectly static. Rather, it fluctuates with a month time scale, within ±1.5°. Across 2.5 million km, this means that the laser could be pointing off about 50000 km, which would not allow any measurement. Hence, an active pointing of laser beams is required. This task will be handled by the MOSAs, the Moving Optical Sub-Assemblies, which will provide high-precision pointing of the laser beams toward the distant spacecraft, moving LISA's 30 cm telescopes accordingly. Telescope movement should be as steady as possible, not to induce additional noise and spoil GW measurement, meeting the requirement of $10\,\mathrm{nrad\,Hz^{-1/2}}$. The relative tilt between the local laser and the incoming one from the distant spacecraft will be measured exploiting once again the laser link, comparing the relative orientation of the incoming beam wavefront and the local reference laser. As shown in Fig. 1.4, particularly in the central panel, the telescopes will be rigidly bound to the optical benches, on which the interferometric measurements will be performed.

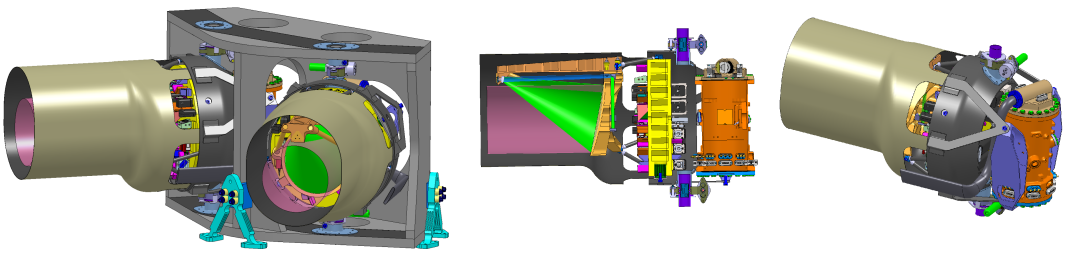


Figure 1.4: View of the LISA scientific payload, as in one of the first designs. In all three images, the telescopes can be seen, as well as the MOSAs, pointing the telescopes to the distant spacecraft. In the central and right pictures, also the GRSs and the optical benches can be seen. Adapted from [6].

Interferometry on LISA (Figs. 1.5, 1.6) will be inherently different from the usual on-ground one. Though the setup might resemble a set of three Michelson interferometers, the reality is substantially different. The emitted laser, a 1064 nm infrared beam with a power of the order of 2 W, reaches the distant spacecraft with no greater power than a few 100 pW, because of aperture loss during its long-arm path. This makes it unfeasible to reflect it for interferometry. Rather, a phase-lock system operates as a transponder, sending back a beam at the full 2 W power, with the same phase as the incoming one. The incoming beam also interferes on the local optical bench with the local laser beam to obtain a phase signal, which ultimately is the measurement carrying GW information. The phasemeter readouts will be then telemetered to Earth, with no further onboard processing.
On ground, phase measurements will be used to synthesize virtual interferometric arms, yielding the measurement of the gravitational strain. The post-processing technique that will be implemented to synthesize the virtual arms is known as Time Delay Interferometry (TDI); the objective of the TDI is to construct virtual interferometric arms with the smallest length difference achievable, which is essential for the mitigation of the impact of laser frequency noise on GW detection.

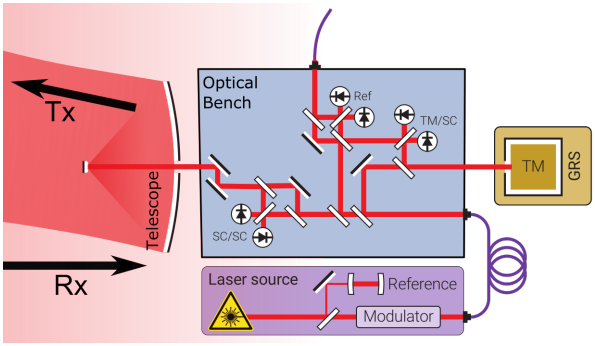


Figure 1.5: Scheme of the interferometric measurements performed within a MOSA, representing the local TM-to-SC measurement and the distant SC-to-SC measurement. Adapted from [8].

From the instrumental point of view, long-arm interferometry must also face another issue. The spacecraft, during the orbital motion, naturally maintain the equilateral triangle configuration only to a certain extent. The nominal 2.5 million km inter-spacecraft distance will be maintained within ± 35 000 km, varying on a yearly basis with peak velocities of the order of 10 m/s. This velocity translates to a Doppler shift of the laser beam: given its frequency of the order of 300 THz, 10 m/s corresponds roughly to a 10 MHz frequency shift.
This issue is overcome by accurately changing the frequency of the outcoming beam, to compensate for the Doppler shift it will experience. Based on a pre-computed frequency plan, the interferometric beatnotes will always be kept between 6 and 25 MHz, allowing for heterodyne interferometry [8].
Lastly, the main focus of this work, the gravitational free-fall reference.
The test masses in LISA will play a central role, being the ultimate geodetic references, complying with the acceleration requirements for GW observation. Along the sensitive

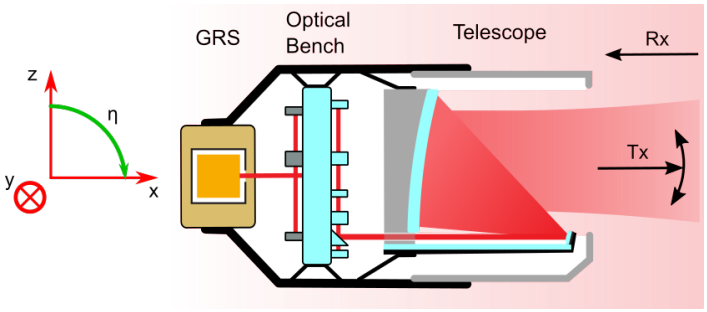


Figure 1.6: Side-view of a MOSA, showing a schematic representation of the GRS, the optical bench, and the telescope. Adapted from [8].

interferometric axes – i.e., the TM-to-TM axes –, TMs will be able to free-float, not constrained by any on-purpose applied force; it will be the spacecraft to follow them in their motion, using in-loop local interferometry. Along the other degrees of freedom, electrostatic forces will be applied, in closed loops with displacement sensing, to ensure stability.
The core system designed to provide reliable geodetic references is known as the GRS, the Gravitational Reference System. That system is designed to control and sense the positions of the TMs, and to provide a stable low-noise environment around the TMs. It is also meant to cage firmly the TMs during launch, to provide a suitable vacuum environment before and after launch, to grab the TMs during commissioning and whenever needed, to provide UV-light discharge, and in general to perform core operations on the TMs, to ensure that they are reliable free-falling references. The GRS is well visible in Fig. 1.4, depicted in orange. The GRS for LISA will extensively inherit technologies from LPF, hence we postpone a more thorough description to Sec. 2.1.1. In general, the GRS is responsible for the acceleration noise performances of the mission, whose assessment and analysis are the main focus of this work.

The current top-level requirements [12] set the following expression as, respectively, the single-TM maximum acceleration noise and the single TM-to-TM laser-link accuracy, in the frequency band 100 µHz–1 Hz:

$$S_g(f) < \left[2.4\,\mathrm{fm\,s^{-2}\,Hz^{-1/2}}\right]^2 \times \left[1+\left(\frac{0.4\,\mathrm{mHz}}{f}\right)^2\right]\left[1+\left(\frac{f}{8\,\mathrm{mHz}}\right)^4\right] \tag{1.3}$$

$$S^{1/2}_{x,\mathrm{TM\text{-}TM}}(f) < 10\,\mathrm{pm\,Hz^{-1/2}}, \tag{1.4}$$

We do not go further deeper into the details of the LISA mission, as it is not needed for the purposes of the present work. In the following, we discuss the LPF mission. We will see that the acceleration performance requirement of LPF (Eq. (2.1)) was looser than that of LISA, as it was intended as a technology-test mission. We will describe the main systems onboard LPF, their properties, the actuation schemes, and the data production pipelines.

2 | LISA Pathfinder

Designed to assess the technological feasibility of the LISA mission, LPF was launched on December 3rd, 2015, from ESA's spaceport in Kourou, French Guiana to the Sun-Earth Lagrange point L1. The mission was not designed to detect gravitational waves, but rather to demonstrate from the technical perspective that rigid bodies can be set in space in nearly-geodesic motion, complying with the requirements for gravitational wave observation. Secondarily in its scope, but not less importantly, LPF served as a reference for high-precision metrology, marking important landmarks for future missions. It represented a milestone in the design of low-noise measurements in space, being a unicum under many scientific and engineering aspects.

The main objective of LPF was to prove that masses can be set in space freely falling, following geodetic trajectories as closely as required to detect gravitational waves. Since the scope of LISA is the detection of $\sim$ mHz and sub-mHz GWs, LPF was designed to work in the frequency band from 1 mHz to 30 mHz, and to achieve a maximum stray acceleration noise of

$$S_{\Delta g}^{1/2}(f) \leq 30\sqrt{1+\left(\frac{f}{3\,\mathrm{mHz}}\right)^4}\ \mathrm{fm\,s^{-2}\,Hz^{-1/2}} \tag{2.1}$$

The main scientific measurement, appearing in Eq. (2.1), was indeed a differential acceleration, between two gold-platinum test masses, which were floating inside the spacecraft without any mechanical contact with its surfaces. In the mission concept, one of the test masses (TM1) served as a free-fall reference, along the TM-to-TM axis. The spacecraft itself was part of the scientific payload of the mission, since it was equipped with actuators, essential for the scientific result itself. The actuators were needed to apply µN forces to the spacecraft, keeping it in the nominal position with respect to TM1, hence counteracting the external

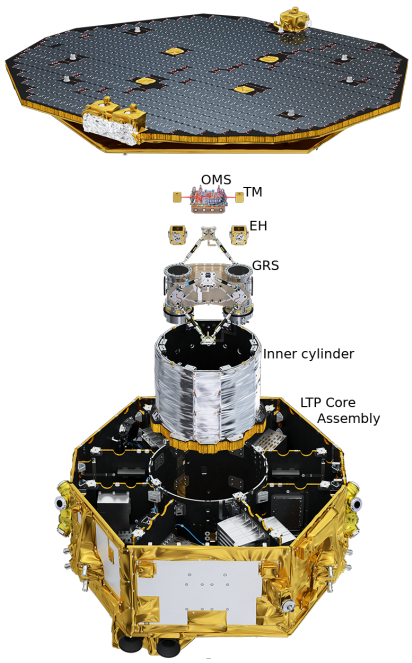


Figure 2.1: Rendering of the LPF spacecraft, representing the main instruments, also described in text. [Credits: ESA/Medialab]

forces with a drag-free scheme. A second test mass, TM2, was needed to perform a differential measurement, with respect to the freely-falling reference represented by TM1. This scheme, known as drag-free control, allowed for a crucial noise reduction, effectively shielding the test masses from external influences.

Figure 2.2 depicts a very schematic representation of the LPF system, with some acronyms and details that will be described later.

The most sensitive measurements were performed with an interferometric system, also visible in figures in the following pages, which allowed to reach the sub-pm-level measurements. LPF hosted the first sub-pm interferometer ever flown in space [13], which performed much better than its requirements. Motion along the remaining degrees of freedom, both linear and angular, was controlled by electrostatic forces, provided by electrodes in close proximities to the test masses. Such electrodes also provided independent capacitive positioning sensing.

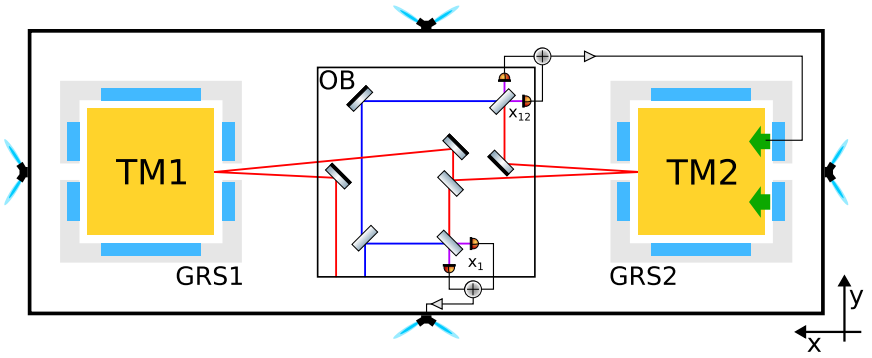


Figure 2.2: Schematic representation of the main LPF concept, the main systems, and loops. The gold squares represent the test masses, caged in their electrode housings, part of the GRS subsystems. The interferometric measurements $x_{1,\mathrm{OMS}}$ and $x_{12,\mathrm{OMS}}$, operated on the optical bench, provide the input signals for the control loops and constitute the basis for the scientific measurements. The green arrows indicate that TM2 is actuated along the TM-to-TM axis. Only test masses and electrodes are represented with real proportions.

This chapter provides an overview of the main instruments flown and operated onboard LPF, with a general description of the most important features, and a focus on the quantities and parameters that will be useful for the analyses proposed in this work.
In the first section of this chapter, 2.1, we describe the LPF experiment, its scientific payload, the main measurement systems; we describe LPF operations, experimental runs, and conditions. In the second section of the chapter, 2.2, we describe the data analysis and post-processing of the telemetered data, the results that were achieved, and give useful data series definitions.

2.1 LISA Pathfinder: mission and experiment

Many systems flew onboard the LPF spacecraft, crucial to the success of the experiments. As already said, LPF was, as LISA will be, a unicum from some points of view: the spacecraft themselves do not only host the payloads, but rather have an

active role in the scientific measurements, following the test masses to reduce the influence on their motion. LPF hosted two different spacecraft thruster mechanisms, constituting two payloads: the LISA Technology Package (LTP) and the Disturbance Reduction System (DRS). The LTP thrust mechanism, which was used for most of the time and which will be employed on LISA, used cold nitrogen, released by nozzles, to transfer momentum. The second one, the DRS, was provided by NASA, and was used in the so-called ST-7 runs. It used colloidal materials, propelled at high velocity, to transfer momentum with high mass efficiency.

The two drag-free designs implemented different control schemes, but shared the same metrology and reference system provided by the LTP, namely, the Optical Metrology System (OMS), and the Gravitational Reference Sensor (GRS) and providing the inertial test mass references.

2.1.1 GRS: gravitational reference sensor

The LTP hosted two gravitational reference sensors (GRSs). The scope of each GRS was ultimately to control actively and to sense the positions of the test masses, located at their centers with no contact. As shown in Fig. 2.3, each GRS contained a TM, and the two were interconnected by the interferometric positioning sensing.

The GRS took care of the inertial free-falling references, i.e., the test masses. First, the systems provided a controlled vacuum environment, being equipped with two separated vacuum chambers, evacuated and sealed on ground before launch, and vented to space once in orbit. It provided the caging, grabbing, and releasing mechanisms (CVM and GPRM), essential for the correct smooth free-fall initiation. Each test mass was also surrounded by an electrode housing (EH), which provided active control of the TM motion along all the non-drag-free degrees of freedom. We usually refer to control forces as to *actuation*. The GRS also provided a static gravitational balance of the spacecraft field, to relieve the actuation from applying strong forces.
The electronics also provided capacitive sensing, relevant wherever optical interferometric sensing was not available.
Lastly, the GRS hosted the TM UV-light system, essential for TM discharging.

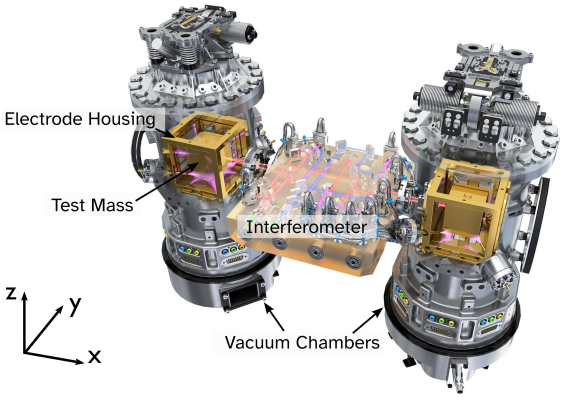


Figure 2.3: Rendering of the LISA Technology Package. The rendering shows the two test masses hosted inside their respective electrode housings (some of the electrodes are not represented), and the vacuum chambers enclosing both test masses and electrode housings. The picture also shows the high-stability optical bench hosting all interferometric readouts, and many other features of the instrument, launch lock, UV-light-based test-mass neutralizer, etc. that are not relevant here.
[Copyright: ESA/ATG medialab]

The test masses

At the core of the LPF experiments, there were the test masses (TM), two gold-platinum cubes acting as free-fall geodesic references, identical to the ones that will be used on LISA. The TMs were nominally located 37.6 cm apart, weighted (1.928 ± 0.001) kg, and were (46.000 ± 0.005) mm wide.

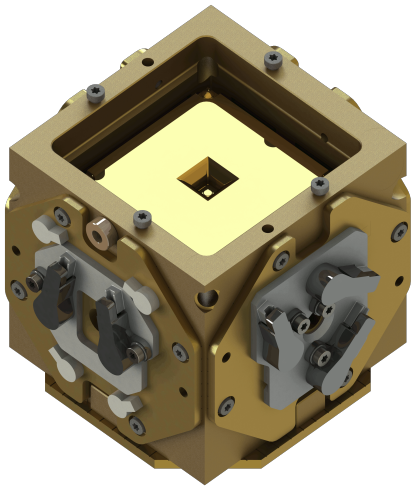


Figure 2.4: Rendering of a gold-platinum test mass flown on LPF, inside its electrode housing. The caging indents can be seen at the TM corners, as well as the central grabbing recess. Courtesy OHB Italia.

The choice of the Au-Pt alloy, 73%Au/27%Pt, was crucial for some reasons. First, it allowed the test masses to have a high density, and a considerable mass despite their small size. Such an alloy also allowed for high electrical conducibility, essential for electrostatic sensing and actuation, and a high thermal conducibility. The surfaces were gold-coated, as they were needed to act as the end mirrors of the interferometric system. The alloy composition also allowed to tune the magnetic properties of the final product: combining gold and platinum, respectively diamagnetic and paramagnetic metals, allowed to achieve a final material with $\chi \sim -2.5\times10^{-5}$, a very low susceptibility, with slightly negative diamagnetic behavior [14]. The magnetic moment of the TMs was also measured before flight [15], resulting to be zero within the instrument error, $|\mu| < 5\,\mathrm{nA\,m^2}$. Indeed, if not taken care of, the TMs motion would be spoilt by interaction with the environmental interplanetary magnetic field, which could potentially be a relevant force source.

On the z faces, the TMs were machined to have small 2 mm-sized spherical indents at the corners, needed for launch caging. These indents were the only non-coated part of the TM, needed to prevent cold welding to the fingers of the CVM mechanism. Additionally, the TMs had quite large ~ 1 cm central recesses on the z faces, to allow mechanical re-grab with the GPRM plungers, which was needed for example during station-keeping maneuvers, or in emergency cases.

The electrode housings

The TMs were surrounded by electrodes, positioned as shown in Fig. 2.5. They were essential for the nominal science-mode functioning of the LPF experiments, providing active control and sensing of the TM motion along all degrees of freedom [16]. The electrodes were enclosed in electrode housings, gold-coated molybdenum structures which hosted both the injection and sensing/actuation electrodes, as shown in the figure. The electrodes as well were made of molybdenum, gold coated on the TM-facing side, and attached to the housings via sapphire insulators.
From the capacitive-sensing point of view, it would be recommended to have small

gaps between the electrodes and the sensed object, as narrow as possible to increase sensitivity [17–19]. Nevertheless, gaps for LPF were chosen to be $d_x = 4.0\,\mathrm{mm}$, $d_y = 2.9\,\mathrm{mm}$, and $d_z = 3.5\,\mathrm{mm}$. Such large gaps require high sensing bias voltages [16], to compensate for the drop in sensitivity due to large gaps. The compensation is however only partial, and one ends ultimately with a worsened sensing resolution (which on LPF was at the $\mathrm{aF\,Hz^{-1/2}}$ level, over a pF nominal static level). This corresponded to a positioning $\mathrm{nm\,Hz^{-1/2}}$ noise, much higher than the $32\,\mathrm{fm\,Hz^{-1/2}}$ reached by the interferometric system. The reason for this choice is force noise reduction. It has indeed been shown [20–22] that a number of effects contribute stronger to the force noise level when the gap is narrow. Such effects include stray patch potentials, and increased Brownian force noise. The chosen gap sizes were eventually selected as a trade-off.

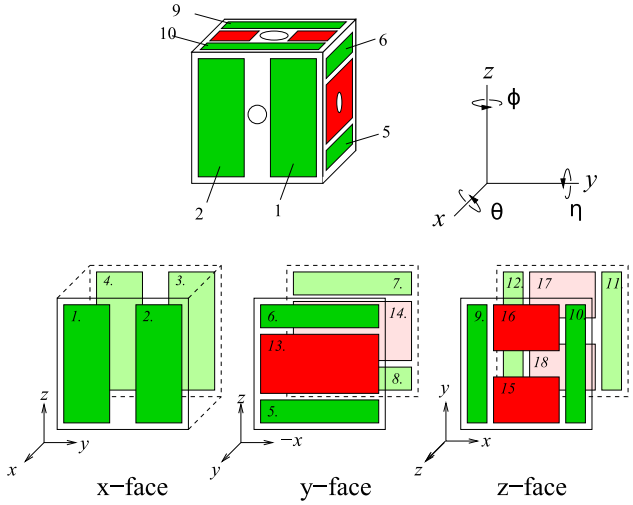


Figure 2.5: LPF electrode configuration. Actuation/sensing electrodes are depicted in green, while the injection ones are shown in red. Electrodes on the x face are used to control the x and ϕ degrees of freedom; electrodes on the y face, to control y and η; electrodes on the z face, to control z and θ.

The electrodes served a number of different purposes:

1. Inject a 100 kHz AC voltage, through electrodes depicted in red in Fig. 2.5, meant to polarize the test masses for positioning sensing.

2. Sense the TM position and orientation along all degrees of freedom, through electrodes depicted in green in Fig. 2.5. Differential measurements between selected electrode pairs were done via capacitive-inductive bridges [16], demodulating the 100 kHz detected signal.

3. Apply audio-frequency actuation signals, meant to apply forces and torques to the TMs, as required by the actuation settings and calculated by the control systems.

4. Apply DC voltages to the electrodes, meant to counteract DC stray potentials appearing because of potential patches on the TM surfaces. The configuration

of this DC compensation was changed sometimes during the mission, as detailed in App. A.

Electrostatic actuation and sensing

As said, electrostatic TM control provided both a capacitive positioning readout and actuation systems [16, 23]. We recall that the TMs did not touch the EHs and thus had no grounding wires, since such mechanical parts would generate a thermal force noise much higher than the LPF requirements. Hence, both sensing and actuation were purely electrostatic, and use the same set of 18 electrodes. Two different frequency bands were used for sensing and actuation, to avoid cross-talks as much as possible. However, in Chapter 6 we describe a low-frequency cross-talk that we detected, between actuation and sensing.

Sensing. As described in [16, 21], sensing operated at 100 kHz frequency: 6 of the 18 electrodes, located centrally on the y and z faces, were used to inject a 0.6 V zero-mean signal. The remaining 12 electrodes, 2 per TM face, were instead used for homodyne sensing of the capacitance between electrodes and the TM, with capacitive-inductive bridges. Ultimately, the device sensed the translational and rotational positions, inferred from the measured capacitances. Electrodes showed a constant sensing precision across all the LPF mission, with best performances between $0.7\,\mathrm{aF\,Hz^{-1/2}}$ and $1.8\,\mathrm{aF\,Hz^{-1/2}}$ at high frequencies (above 1 mHz, where readout noise dominates). Such sensitivities correspond to $2.4\,\mathrm{nm\,Hz^{-1/2}}$ and $170\,\mathrm{nrad\,Hz^{-1/2}}$, respectively for translational and rotational measurements.

Figure 2.6 shows a block diagram of the capacitive sensing and actuation system, for a pair of electrodes.

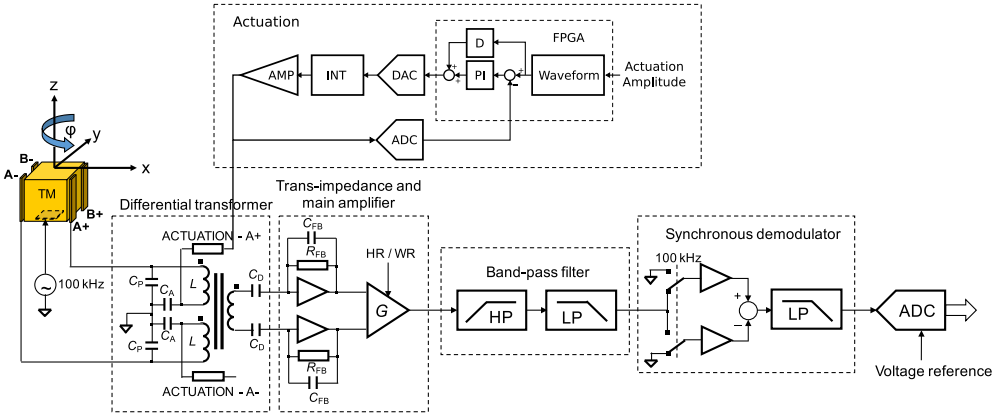


Figure 2.6: Block diagram representing the capacitive sensing and actuation scheme, for a pair of opposing electrodes (A, in the figure), sensing along the x-ϕ directions. Specifically, readouts from the A and B bridges in the figure were averaged and subtracted for (respectively) x and ϕ measurements. Figure adapted from [16] and [23].

Actuation and authority. Differently from sensing, actuation operated in the audio-frequency band, through the application of amplitude-modulated sinusoidal

voltages in the 60-270 Hz band. The choice of high-frequency actuation was driven by sensitivity needs, not to convert the low-frequency voltage noise into a large force. Moreover, different carrier frequencies were used to control different degrees of freedom. Amplitudes were commanded based on the force required by the DFACS (the control system, as described later), to apply the necessary forces and torques. Since the force F applied by a pair of electrodes is proportional to the squared voltage difference ΔV^2, the application of sinusoidal voltages results in the application of forces at twice the carrier frequency and at the carrier amplitude-modulation frequency. Hence, actuation resulted in an effective force following the carrier modulation, with amplitude proportional to the square carrier amplitude[1], $F \propto V_c^2$.

Forces and torques on the TMs were obtained with combinations of voltage waveforms, applied to the corresponding electrodes. As represented in Fig. 2.6, waveforms were computed, digitally synthesized, converted, and then applied to electrodes.

The actuation scheme was designed to apply a controlled and steady force gradient on the TMs[2]. Indeed, it is inevitable that the application of force-inducing voltages is also associated with the generation of a force gradient, in the surroundings of the TMs, translating into forces as the TMs move. Force gradients, as will be better explained in the data processing section, 2.2, on page 31, are listed as noise sources on LPF. Gradients create elastic-like couplings between the test mass position and its acceleration. Although inevitable, it is undesirable that the gradient stiffness depends on the instantaneous commanded force, introducing time-varying force gradients. If it were the case, the data post-processing would require disentangling the effects, computing at any moment the local gradient. To avoid this dependence, LPF exploited an algorithm to apply forces with a constant stiffness [23, 24]: the force gradient did not depend on the applied forces, but rather on the maximum exertable force.

The maximum exertable forces (and torques) are known as *authorities*. During each LPF experimental run, authority was a fixed parameter. Fixing the authority also fixed the impact of actuation on the low-frequency noise, fixing the actuation force noise. Hence, authority was chosen as low as possible on the basis of the highest required force, to minimize the impact of actuation on the force noise itself. We point to App. A, where we list all the actuation configurations used on LPF. The configuration acronyms are also used in Table A.1, listing all the LPF runs used in this work and the employed settings.

DC voltage compensation. Electrodes were also used to apply static DC voltages to the TMs, to counteract spurious voltages. Before launch, it was indeed noted [25, 26] that the LPF setup could be subject to stray voltages, arising from the presence of spatially varying DC surface potentials, known as patch fields, potentially generating force on the TMs. In a simplified model, the potential is uniform over each electrode, which can therefore be assigned a single value. The way those stray voltages would generate in-band forces is the following:
Suppose that the four TM-facing electrodes on the x faces (follow the electrode

[1]The mean squared carrier amplitude, averaged over a carrier period.
[2]Force and torques. For clarity, we just refer to forces.

numbering in Fig. 2.5) are affected by stray DC potentials δV_i, then a force is induced on the TM,

$$F_x = -\frac{q}{C_T}\sum_i \delta V_i \frac{\partial C_x}{\partial x}, \tag{2.2}$$

where q is the total charge on the TM, C_T the total capacitance of the TM to all the surrounding surfaces, and $\partial C_x/\partial x$ the partial derivative of the TM capacitance to an x electrode, with respect to motion along x. This equation can be further simplified by defining the effective potential $\Delta_{x,j} = (\delta V_{\mathrm{x1},j} + \delta V_{\mathrm{x2},j} - \delta V_{\mathrm{x3},j} - \delta V_{\mathrm{x4},j})$, where j represents GRS 1 and 2. The quantity $\Delta_{x,j}$ can be seen as an effective single-electrode potential that would yield an equivalent force along x. Analogously, one can define $\Delta_{\phi,j} = (\delta V_{\mathrm{x1},j} - \delta V_{\mathrm{x2},j} - \delta V_{\mathrm{x3},j} + \delta V_{\mathrm{x4},j})$, which is the analogous quantity for torque along ϕ. Overall, the expressions for static forces and torques read:

$$F_x(q) = -\frac{q}{C_T}\left|\frac{\partial C_x}{\partial x}\right|\Delta_x, \qquad N_\phi(q) = -\frac{q}{C_T}\left|\frac{\partial C_x}{\partial \phi}\right|\Delta_\phi, \tag{2.3}$$

here, $\partial C_x/\partial x$ and $\partial C_x/\partial \phi$ are the partial derivatives of single electrode capacitances with respect to motion along x or ϕ. Any fluctuation in q, hence, induces stray noisy forces on the TMs. A strategy to reduce this source of noise was devised as counteracting the stray potentials Δ_x and Δ_ϕ, obtained by measuring their value and applying counteracting voltages to the electrodes, in a combination that generated both the required Δ_x and Δ_ϕ. Such DC voltages were of the order of $\sim 10\,\mathrm{mV}$. The final DC voltages, after compensation, were estimated as $< 5\,\mathrm{mV}$ [2]. In Table A.5, we list all the used voltage configurations.
We will also note later that imperfect digitization of the AC actuation waveforms led to fluctuating quasi-DC voltages, representing a force source, analogous to Eq. (2.3).

GRS: gravitational compensation and stiffnesses

We have described the LPF system as aimed at following and measuring the motion of the two test masses in their free fall along geodesics, following TM1 and actuating TM2. Nevertheless, the presence of the spacecraft itself changes the geodesic lines themselves, attracting the TMs and generating a net force on them. The gravitational attraction between the spacecraft and the TMs manifests as steady DC forces and torques, that require force compensation to maintain the TM nominal position. Because of the drag-free design (see Sec. 2.1.3), actuators on the spacecraft maintain TM1 in its nominal position within it. Thus, any gravitational imbalance would require DC suspension forces to be applied electrostatically to TM2. As

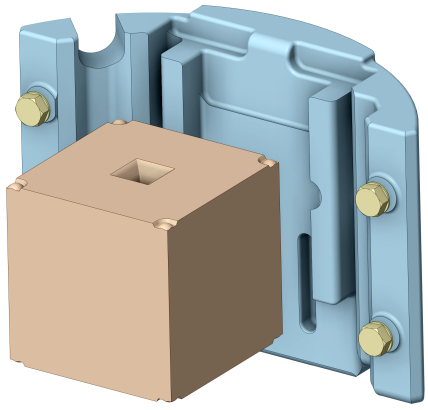


Figure 2.7: Rendering of an Internal Balance Mass (IBM) in GRS1, to scale. Details of the LPF structure, mechanisms, and components have been omitted. Courtesy OHB Italia.

said, the necessity of strong DC forces would
require high actuation forces and high authorities, hence leading to increased noise in the science measurements. It was shown [27] that if the difference in gravitational force per unit mass exceeded $2.5\,\mathrm{nm\,s^{-2}}$, actuation noise would already reach the mission requirements. Hence, strategies were applied to reduce as much as possible the gravitational imbalance, in order to require just a tiny amount of electrostatic forces to compensate for the residual imbalance.
Incidentally, we note that the gravitational forces also cause a force gradient at the TM locations, causing on each TM a force per unit mass that can be effectively modeled as

$$g_{\mathrm{grad},i} = \omega_i^2 x_i,$$

where x_i represents the off-nominal displacement of TMi, and ω_i^2 is an equivalent stiffness, given by the gravitational gradient at the TM location.

To reduce the influence of the spacecraft on the differential test mass motion, the gravitational field was compensated, with high-density masses positioned in strategic locations [27]. Balance masses were placed on the spacecraft itself, and on the external walls of the vacuum chambers. The more critical item for gravitational balance compensation was however the Internal Balance Mass (IBM), placed inside the vacuum chamber of each GRS, as shown in Fig. 2.7. The IBMs must be located in very close proximity to the TMs, to provide the required compensation of gravitational forces and torques. Based on calculations of the overall gravitational field, the IBM was designed and finely machined to compensate an imbalance of about $50\,\mathrm{nm\,s^{-2}}$, down to a $10\,\mathrm{pm\,s^{-2}}$ accuracy, and to keep the gravitational gradients within requirements.

The IBMs were two $\sim 2\,\mathrm{kg}$ uncoated pieces of a sintered tungsten-copper alloy (W/Cu 90/10 for IBM2 and W/Ni-Cu 95/5 for IBM1). The IBMs were designed to "wrap" the TMs from the external side of the line joining the two TMs, a few centimeters away from each TM.

GRS: vacuum handling and venting

Vacuum handling was a sensitive issue to the success of the mission. Indeed, the residual gas surrounding the TMs is itself a source of force noise, and it is well visible in the LPF data series. Pressure evolution is also a central topic of this work; its impact on the mission results, its evolution, and its possible sources are discussed in Chapter 5. As described in [22, 28], the viscous damping caused by the gas surrounding the masses poses limitations to the noise performance, appearing as a frequency-independent white force noise. This result is a direct consequence of the fluctuation-dissipation theorem [29], which states that the presence of a force-dissipation mechanism in a system is always associated with a force noise. In this case of Brownian motion, the noise is frequency-independent, and its amplitude depends on temperature, pressure, the mass of the emitted gas species, and the geometry of the system.

At cryogenic temperatures, this noise source could be negligible. But LPF did not work at low temperatures. In ordinary runs, LPF operated in a temperature range between 284 K and 295 K, hence the noise induced by Brownian motion played a relevant role. The same references cited before [22, 28] show that the white noise

level is proportional both to absolute temperature and to pressure. With the same geometry as LPF's (and LISA's) TM and EH configuration, it was measured that the conversion coefficient between pressure and single-TM acceleration, assuming that the emitted gas is water, is $S_g/P_{H_2O} = (1.3\,\mathrm{fm\,s^{-2}\,Hz^{-1/2}})^2/\mu\mathrm{Pa}$. A pressure greater than 230 µPa would have on LPF's performance an impact as high as its requirement; LPF's pre-flight goal was 10 µPa, and 1 µPa was reached at the end of the mission. The corresponding threshold for LISA would be 3.4 µPa; an envisioned goal for LISA is 1 µPa [12].

The simple relation between acceleration PSD and pressure also serves as an indirect measurement of the static pressure level. Indeed LPF was not equipped, and LISA will not be equipped, with pressure gauges in the surroundings of the TMs, because of their impact on both volume and noise performance.

LPF profited from the space environment, using venting-to-space ducts as vacuum pumps. As shown in Fig. 2.8, each GRS was enclosed in a Ti alloy vacuum chamber, called Vacuum Enclosure (VE), containing all the mechanisms and parts that needed to be in close contact to the TM, but being isolated from the outer outgassing environment. Lasers for interferometry could pass through optical windows. The venting ducts, combined with the GRS geometry, allowed an effective EH air(water) pumping of about 16(19) L/s (OHB Italia estimates).

To reach the desired pressure goal in the short time of the LPF mission, however, venting to space was not enough. Thus, the VE was pre-evacuated and pre-baked on ground. Approximately one year before launch, the VE flight models were baked for about 24 hours, at about 115 °C, after which the inner pressure reached a few post-cooldown µPa. The VEs were subsequently sealed with a valve with Viton gaskets, and opened only once in orbit. Incidentally, we note that the bakeout could not last longer or be performed at a higher temperature,

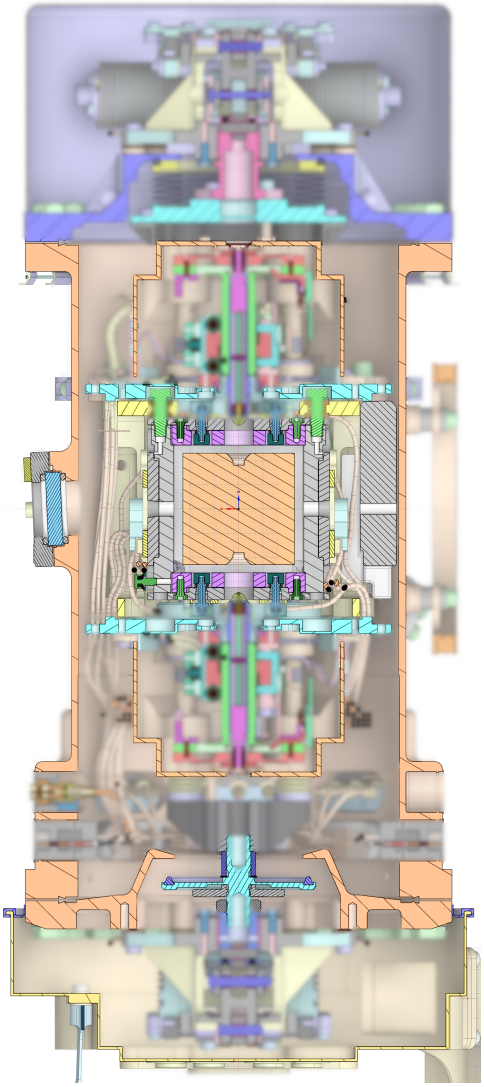


Figure 2.8: Section on the x-z plane of the GRS. The test mass is visible at the center, with the grabbing recesses. y-face electrodes are visible; no electrodes on the x face are visible. On the right side, the IBM. In the lower part, the venting valve, and, at the bottom-right, the venting duct aperture. Non-essential components are blurred. [Courtesy OHB Italia]

because of constraints posed by the emitting surfaces of the charge management system.

The venting mechanism was part of the Caging and Venting Mechanism (CVM), which provided both the vacuum sealing/opening of the inner GRS, and the TM launch-lock mechanisms. Before launch, it guaranteed both the sealing of the VE, and the TM launch-lock, caging it with eight fingers and applying a total of 1200 N, preventing any damage from launch vibrations. Once on orbit, on DAL 62.4, the CVMs unlocked the TMs and vented the VE to space (see Fig. 2.15 on page 29). When opened on orbit, inner pressure was crudely estimated to be of the order of 1 mbar. At the beginning of the LPF science operations, about 40 days after venting, pressure had decreased to about 10 µPa, corresponding to a (water) outgassing of about 200 µPa L/s – at the end of the mission, it reached about 20 µPa L/s.

GRS: TM charging and management

As described by Eq. (2.3), electrostatics induce forces on the TMs, deviating them from their geodesics. Two disturbances of electrical origin have been identified as causing force noise on the TMs: fluctuations in the test mass voltage, and fluctuations in the electrode voltages, following from Eq. (2.3). The first one is inevitably connected to test mass charging in time, caused by the space environment. In both LPF and LISA, the test masses accumulate charge in time: the interaction with cosmic rays, both through direct interactions with the TMs and through secondary emission within the EHs, leads to a net test mass charging [30–32]. Indeed, the TMs are not grounded, and any charge deposited on the TMs will not leave it on its own. The work function of gold surfaces is about 5 eV, which is much higher than thermal fluctuations at LPF temperatures. The charging rate was measured [26], as a net positive rate of $22.9\,\mathrm{e\,s^{-1}}$ on TM1 and $24.5\,\mathrm{e\,s^{-1}}$ on TM2, with variations of the order of $\pm 2\,\mathrm{e\,s^{-1}}$. Assuming a run lasting 15 days, the total accumulated charge would be of the order of pC, corresponding to a ~ 150 mV TM voltage.

The net TM charge induces force noise, enhancing the random fluctuations of the electrode voltages. However, TM charging contributes through another mechanism to the force noise, which is equivalent to shot noise: TM charge is contributed by different species of charged particles, with either positive or negative charges. Those charges cancel out as regards the net charging rate. Nevertheless, hitting the TMs, they generate charging shot noise, which is proportional to the square total charge of each charging event, independently of its sign. This effect contributes through the *effective* charging rate, rather than the *net* charging rate. On LPF, it was measured $\lambda_{\mathrm{eff},1} = (1060 \pm 90)\,\mathrm{e\,s^{-1}}$ and $\lambda_{\mathrm{eff},2} = (1360 \pm 130)\,\mathrm{e\,s^{-1}}$ [26].

As already discussed, a first strategy for reducing the contribution from charge fluctuations is to compensate the stray DC voltages as much as possible. Another strategy implemented to reduce TM continuous charging is direct TM discharge. TM charge was controlled using UV lamps, which shone light on both the TMs and the surrounding EHs: light acts extracting electrons from the gold surfaces via photoelectric effect, provided that the photon energy is higher than the extraction work function. A rendering of the UV photoemission device is visible in Fig. 2.3.

Before each science run, the Charge Management System (CMS) was enabled, to adjust TM charge and prevent it to increase too much. If possible, looking forward to the end of each run, TM charge was set negative before the run start, to keep it as close to zero as possible during the run itself.

The estimates of noise from these sources are given in Chapter 6.

2.1.2 Interferometry, Optical Metrology System

The position of the TMs along the sensitive degrees of freedom was interferometrically sensed, as schematically shown in Fig. 2.10. The interferometric system [13, 33, 34], part of the Optical Metrology System (OMS), consisted of four interferometers IFO, residing on an Optical Bench (OB). The interferometers were used to sense:

- $x_{1,\mathrm{OMS}}$, the displacement between TM1 and the OB. Given that the OB can be effectively considered as rigidly bound to the spacecraft, $x_{1,\mathrm{OMS}}$ is a measurement of the TM-to-SC distance. The readout of this interferometer was the driving signal of the drag-free control loop, driving the µN forces to be applied to the spacecraft along x.

- $\Delta x_{12,\mathrm{OMS}}$, the distance between TM1 and the TM2. The readout of this interferometer was the driver of the suspension loop, driving the fN electrostatic forces to be applied to TM2 along x.

- $\eta_{1,\mathrm{OMS}}$, $\phi_{1,\mathrm{OMS}}$, the angular orientation about the y and z axes of TM1, with respect to the optical bench. The use of quadrant photodiodes (QPD) allowed angular measurements via differential wavefront sensing technique [33].

- $\eta_{12,\mathrm{OMS}}$, $\phi_{12,\mathrm{OMS}}$, the angular orientation about the y and z axes of TM2, with respect to TM1. This measurement allowed to reconstruct $\eta_{2,\mathrm{OMS}}$, $\phi_{2,\mathrm{OMS}}$.

- The remaining six degrees of freedom were not interferometrically sensed. Rather, only the capacitive sensing provided by the GRS was available for those degrees of freedom.

The IFO system consisted of four heterodyne interferometers, operating simultaneously with a shared laser source, following different optical paths as shown in Fig. 2.10. The laser source was an Nd:YAG source providing nominally 35 mW of light at a wavelength of 1064 nm. Nominally, 2.4 mW impinged on TM1 and 1.2 mW impinged on TM2. The laser source was split into two beams, namely a *measurement* beam and a *reference* beam; the reference one, depicted in blue in Fig. 2.10, completely resided on the OB. Hence, its pathlength was artificially increased by about 38 cm, that is the inter-TM distance, to equalize the two pathlengths. The operating principle of the interferometers, operating in

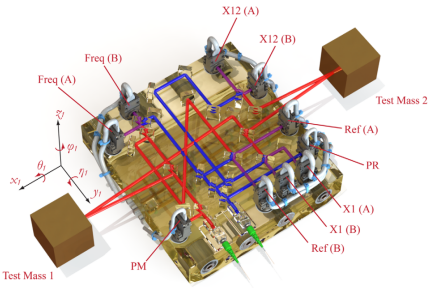


Figure 2.9: Rendering of the LPF optical bench. Image from [13], see Fig. 2.10 for details.

heterodyne mode, required frequency shifts, different in each beam. Acousto-optic modulators (AOM) allowed to shift the beam frequencies by about $f_1 \sim f_2 \sim 80\,\mathrm{MHz}$, with a frequency difference of $f_1 - f_2 = 1\,\mathrm{kHz}$, whose measurement allowed heterodyne interferometry.

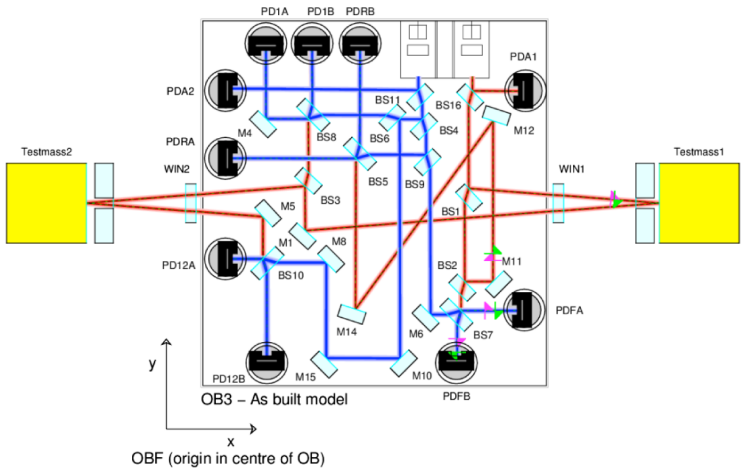


Figure 2.10: Scheme of the OMS system flown onboard LPF, representing *a)* the OB, with all its optical components, mirrors, beam-splitter and photodiodes, *b)* the two laser beams and their optical paths, *c)* the test masses, acting as end mirrors of the measurement beam. Adapted from [33].

We refer to the four interferometers as:

- x1, measuring $x_{1,\mathrm{OMS}}$, $\eta_{1,\mathrm{OMS}}$, and $\phi_{1,\mathrm{OMS}}$, as previously defined.
- x12, measuring $\Delta x_{12,\mathrm{OMS}}$, $\eta_{12,\mathrm{OMS}}$, and $\phi_{12,\mathrm{OMS}}$, as previously defined.
- F, with an optical path residing entirely on the OB, used to sense laser frequency fluctuations and to drive a digital control loop aimed at stabilizing it [34].
- R, with an optical path residing entirely on the OB, sensing common-mode noise shared by all the IFOs. The R readout was digitally subtracted from the readout of all the other interferometers. We note that this subtraction was performed at a rate of 100 Hz, whereas telemetries were downlinked at 10 Hz after subsampling.

The power measured at the photodiodes was also used to drive a power-stabilization loop, aimed at reducing power fluctuation. As we will see later, it is essential to reduce power fluctuation, as the laser radiation pressure on the TMs induces a net differential force on them.
To ensure optimal pathlength stability, the OB was a $(20.0 \times 21.2 \times 4.5)\,\mathrm{cm}^3$ baseplate entirely made of Zerodur$^{\mathrm{TM}}$. The optical components, namely mirrors, beam-splitters, were made of the same material, hydroxide-catalysis bonded [13, 35] to avoid mechanical stresses. This composition allowed high thermal stability and a low thermal coefficient, of order $5 \times 10^{-8}\,\mathrm{K}^{-1}$.

The interferometer system worked much better than expected [1, 2, 13, 33], reaching a much lower noise than requirements. Before flight, requirements for the x1 and x12 were set at linear $9\,\mathrm{pm\,Hz^{-1/2}}$ and angular $20\,\mathrm{nrad\,Hz^{-1/2}}$, at frequencies above $\sim 10\,\mathrm{mHz}$. The measured in-flight performances[3] show that the interferometers performed at a level as low as $32.0^{+2.4}_{-1.7}\,\mathrm{fm\,Hz^{-1/2}}$ for linear motion, and $100\,\mathrm{prad\,Hz^{-1/2}}$ for angular motion. This result is shown in Fig. 2.11, adapted from [13].

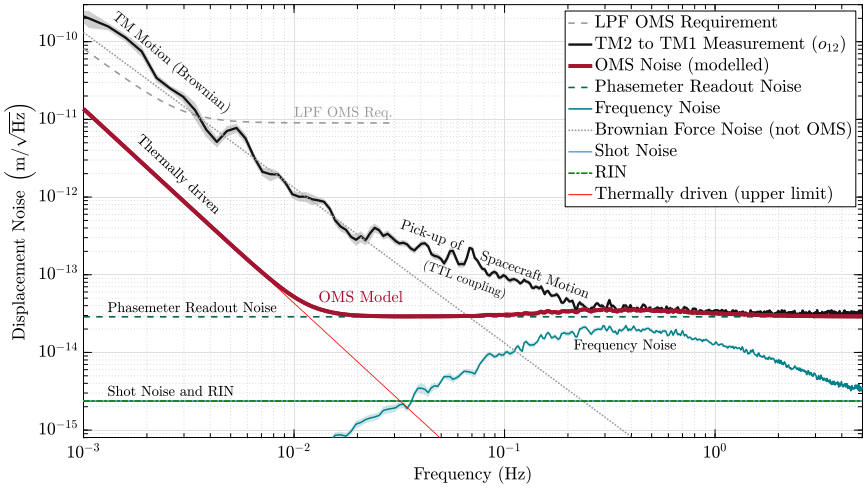


Figure 2.11: In-flight noise performance of the X12 interferometer, compared to pre-flight requirements, and broken down to its main components. Above 10 mHz, the IFO readout noise dominates, at the remarkably low level of $32.0^{+2.4}_{-1.7}\,\mathrm{fm\,Hz^{-1/2}}$. Adapted from [13].

Alignment. The test masses, as end mirrors of the interferometric measurement, were aligned three times during the mission, to ensure optimal laser reflection. Moreover, TM misalignment may generate spurious leakage of jitter along other degrees of freedom into the sensitive measurement. Angular jitter around an off-centered axis, as well as lateral jitter if the TMs are tilted, translates into an effective laser "tilt-to-length" pathlength variation. Mechanisms leading to such coupling are subject to thorough analyses, in view of the LISA mission [36–39]. Regarding LPF, cross-talks were more efficiently removed in post-processing (Sec. 2.2.2), rather than with TM alignments. However, attempts were done during operations to reduce the real cross-talk, aligning the TMs according to the measured coupling coefficients. Table A.4 summarizes the alignment settings.

2.1.3 DFACS: drag-free scheme and attitude control

The DFACS (Drag-Free and Attitude Control System) was the system dedicated to controlling the system dynamics, namely the dynamics of the spacecraft, TM1, and TM2 [40]. It operated to *a)* maintain the spacecraft attitude, *b)* drive the spacecraft to follow TM1 along the TM-to-TM axis, enabling drag-free control, *c)* control TM1 and TM2 with electrostatic forces and torques along the y, z, and angular degrees of

[3] Performance refers to the x12 IFO, as x1 is the in-loop measurement driving the drag-free control, hence picking up the intrinsic spacecraft movement noise. However, the on-ground pre-flight measurements were comparable to the x12 performance [33].

freedom, *d)* control TM2 with electrostatic forces along the x axis, to the reach the differential motion target.

In detail, the DFACS operated to:

1. Control the spacecraft attitude, to keep the solar panels oriented towards the Sun. The loop was controlled by a set of star trackers, driving the spacecraft's attitude with respect to the sky position of a set of stars, and controlling its rotation. However, even at low frequencies below 1 mHz, electronic star tracker noise was the main source of attitude noise, causing an undesired $\sim 10\,\mu\mathrm{rad\,Hz}^{-1/2}$ jitter. Later, we will see that this jitter makes the LPF reference frame non-inertial, coupling to the TM differential motion through a spurious tangential "Euler" force.
   Incidentally, we note that this effect will not be present in LISA. The spacecraft's attitude will not be driven by star trackers but by differential wavefront sensing, comparing the orientation of the local laser beam and the incoming distant one. This will enable a much more sensitive angular control, of the order of $10\,\mathrm{nrad\,Hz}^{-1/2}$ [12].

2. Control the spacecraft along the x direction, based on the x-axis position of TM1 with respect to the spacecraft, i.e. based on the readout of the $x_{1,\mathrm{OMS}}$ interferometer. This loop, known as the drag-free loop, allowed for an efficient disentanglement of the differential TM motion from that of the spacecraft. The rejection ratio of the spacecraft motion from the differential signal was [24] $|\delta^{-1}_{\mathrm{ifo},1}| = (4.5 \pm 0.4) \times 10^4$. This allowed not only to reject the spacecraft motion stationary noise (solar radiation pressure, microthruster noise), but also to isolate the differential TM motion from transient events, such as micrometeoroid impacts [41].

3. Control TM1 and TM2 along the y and z translational directions, applying electrostatic forces to the TMs. The common-mode TM motion is however controlled in drag-free mode, commanding forces to the microthrusters to follow the spacecraft along the common-mode translational and angular TM positions.

4. Control TM1, and TM2 along the angular η and ϕ directions, to keep alignment with respect to the electrodes, applying electrostatic torques. These degrees of freedom are tracked with the differential wavefront sensing of the x1 and x12 interferometers.

5. Control TM1, and TM2 along the angular θ direction, to keep alignment with respect to the electrodes, applying electrostatic torques. This degree of freedom is tracked by the GRS capacitive sensors.

6. Control TM2 along the x direction, based on the x12 interferometric measurement. The control loop controlling this degree of freedom is known as the suspension loop. Given the system geometry, it is not possible to follow both TM1 and TM2 in drag-free mode. Hence, TM1 is chosen as free-falling reference, with the spacecraft controlled to follow it. TM2 is instead actuated in the x direction, keeping its low-frequency position constant with respect to TM1, and,

consequently, to the spacecraft. In Fig. 2.16 on page 34, we show the scheme of the suspension loop, when describing the system dynamics.

2.1.4 S/C control and cold gas thrust

During science operations, the DFACS controlled the drag-free motion of TM1 along the sensitive axis, calculating the necessary acceleration to follow it in its motion, and commanding the force to be applied to the spacecraft itself. During LTP operations, on which we focus in this work, this task was achieved by propelling the spacecraft via two sets of six cold gas thrusters, fed with cold (room temperature $\sim 295\,\mathrm{K}$) nitrogen [42]. These thrusters were also used during station-keeping operations, for orbit adjustments. The two sets, with six thrusters each, provided redundancy and were used one at a time.

Thrusters, located on the outer part of the satellite, applied the required forces and torques, providing a thrust of order µN. Such thrust accelerated the $\sim 500\,\mathrm{kg}$ spacecraft, at the required $\sim \mathrm{nm\,s^{-2}\,Hz^{-1/2}}$ level, guaranteeing the drag-free motion and high rejection of the spacecraft motion from the differential motion. It also guaranteed effective shielding from external events, such as micrometeoroid impacts [41].

Thrusters accelerated the spacecraft in the X+, X-, Y+, Y-, Z+ directions. No additional acceleration was required opposite to the sun-pointing direction, since the solar radiation pressure to the solar panels already provided a 25 µN thrust. With an estimated noise level of $0.4\,\mathrm{\mu N\,Hz^{-1/2}}$ at 1 mHz, thruster noise represented however the dominant source of disturbance noise on the spacecraft, about one order of magnitude above the solar radiation pressure.

About 10 kg of cold nitrogen were kept in four 8.3 L tanks, located as shown in Fig. 2.12 on the right panel, which stored it at a 292 bar pressure. Tanks 3 and 4 were short-circuited, and were effectively assumable as a single tank, which we refer to as tank 3. Only one tank was used at a time, feeding only one of the two thruster sets (A and B). Tank/thruster configurations in the noise-only runs are summarized in App. A.
On the left panel of Fig. 2.12, we show a schematized view of the gas-feeding system: high-pressure nitrogen flowed through pressure regulators, meant to reduce its pressure from about its storage 292 bar to 1 bar, collecting it in a plenum before delivering to the end thrusters. Then, it flowed through valves to the A or B thruster sides, as needed. Pressure in the high-pressure tanks was telemetered, but its readout noise was pretty high to allow precise measurements.

Gas emission from thrusters was controlled with mass flowmeter sensors and in-loop piezo valves, as shown in Fig. 2.13.

Based on the measured $x_{1,\mathrm{OMS}}$, and on the forces/torques commanded by the DFACS, the thrusters' electronics (MPE in Fig. 2.13) computed the required thrust to be provided by each thruster. Then, it commanded a certain mass flow to each valve, based on the conversion factor between the mass flow and applied thrust.

In the following analysis chapters, we will be interested in the fuel mass depletion from the fuel tanks. Here we note that for the thrust loop to work properly it is not

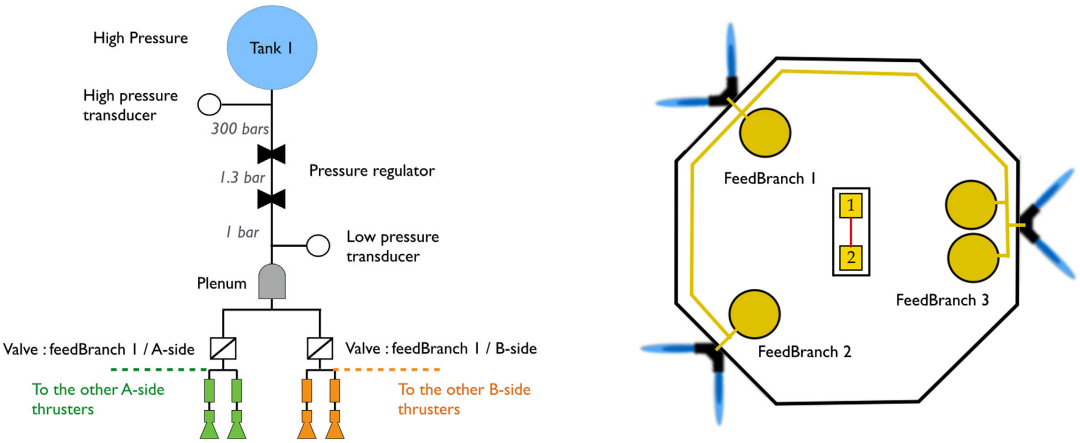


Figure 2.12: (left) Cold nitrogen feed-line, from the high-pressure tank to the feedbranch actuator valve. Tanks were used to feed two redundant, independent thruster sets, named A and B. (right) Position of the three fuel tanks within the spacecraft. Figures from [42].

needed a precise estimate of the calibration factor mass-flow-to-thrust, nor of the valve voltage-to-thrust calibration factor. Hence, the estimate of the tank depletion time series can also be affected by calibration uncertainty, in addition to the intrinsic flowmeter noise. As noted in [42], thrusters were calibrated during station-keeping maneuvers, and were estimated to be precise to better than 10%.

The reason we will be interested in fuel depletion for noise performance, is that the mass loss from fuel tanks can be well identified in the differential acceleration time series. Indeed, as fuel is lost from the tanks, the unequal gravitational attraction between the tanks and the test masses causes the onset of a gravitational force, drifting in time with a constant slope. When in use during LTP operations, thrusters provided a quite constant thrust DC level, using about 10 g of mass per day, which correspond to a force drift $\sim 10\,\mathrm{am\,s^{-3}}$. As shown in Fig. 2.14, the calculated drift, accumulating over time, was responsible for a maximum $\sim 40\,\mathrm{pm\,s^{-2}}$ static drift, which was well visible in the Δg time series. The drift sign depends on the used tank: usage of tanks 1 and 2 mimicks TM attraction (positive Δg, as defined later), due to the loss of gravitational pull between tanks and TMs, while usage of tank 3 mimicks TM repulsion (negative Δg).

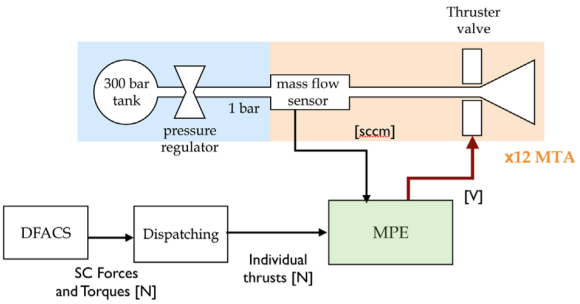


Figure 2.13: Gas propulsion loop. MPE stands for micro-propulsion electronics, controlling the voltage of the thruster valve based on the required thrust. Figure from [42].

Actuation could accommodate for forces of such a small amplitude; however, the tank-usage configuration was changed from time to time, to keep the total differential force as low as possible.

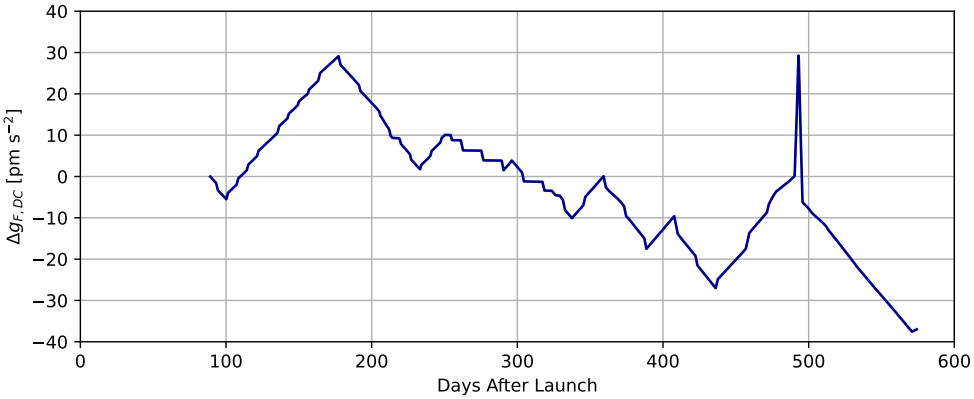


Figure 2.14: Static Δg drift during the LPF mission, caused by the mass depletion from fuel tanks, amounting to $\sim 10\,\mathrm{g/d}$. Positive and negative slopes correspond to the tank in use: usage of tanks 1 and 2 causes TM attraction (positive Δg), while usage of tank 3 causes TM repulsion (negative Δg). Feed branch usage in noise runs is summarized in Table A.2.

2.1.5 LPF orbits

Because of the stringent mission requirement in Eq. (2.1), and envisioning the even more stringent requirement for LISA (Eq. (1.3)), LPF could not operate in a low Earth orbit (LEO), but it needed to reach the Lagrange point L1, at 1.5×10^6 km from Earth, towards the Sun. Indeed, on an LEO, the gravitational gradient ($\sim$ µN/m) would prevent the mission to comply with the low-noise requirements[4]. This point did not represent a limitation for other missions with low-acceleration requirements, such as MICROSCOPE, GOCE, GRACE, GRACE-Fo, but the requirements for LPF were unprecedentedly low[5]. The choice of L1, instead of other Lagrange points, was motivated by operability and communication reasons. Orbiting around L1, which is not a gravitationally stable equilibrium point, only required a few station-keeping maneuvers, during which science measurements were interrupted. Moreover, this orbit allowed the spacecraft to be constantly exposed to sunlight, thermally stable, and able to communicate with Earth.

2.1.6 LPF mission operations

LPF flew between December 3rd, 2015, 04:04:00 UTC, and July 18th, 2017, 19:57:00 UTC, nearly 600 days after launch. In Figure 2.15, we provide a scheme of LPF operations, with details concerning the analyses proposed in this work.

Commissioning started during the journey to L1, with the enabling of the systems,

[4]As noted in the previous section, a high DC force requires a high actuation authority, leading to increased actuation noise. In addition to this, the gravitational gradient in low-Earth orbits is not constant. Because of variations in the mass and density below, the local gradient varies over time as the spacecraft orbits the Earth.

[5]Incidentally, some of the missions here cited, namely, GOCE, GRACE, and GRACE-Fo, were made exactly for geodesy measurements, hence requiring to fly on LEOs. However, other missions' acceleration requirements were not as strict as LPF's and LISA's. The lowest differential accelerations in MICROSCOPE, GOCE, and GRACE were respectively of order $10^{-12}\,\mathrm{m\,s^{-2}\,Hz^{-1/2}}$, $10^{-11}\,\mathrm{m\,s^{-2}\,Hz^{-1/2}}$, $10^{-9}\,\mathrm{m\,s^{-2}\,Hz^{-1/2}}$, to be compared to $10^{-15}\,\mathrm{m\,s^{-2}\,Hz^{-1/2}}$ achieved on LPF [43].

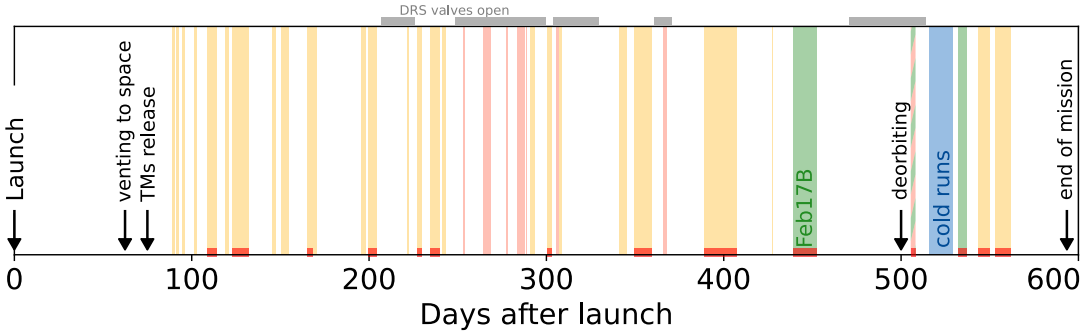


Figure 2.15: Scheme of the LPF runs analyzed in this work, also listed in App. A with more details. Yellow timespans show LTP runs, performed at 22 °C nominal temperature. Green timespans show LTP runs, performed at 11 °C nominal temperature during the mission extension phase. Blue timespans show LTP runs, performed at less than 2 °C. Red timespans show DRS runs, performed at 22 °C nominal temperature. The bi-color span represents a DRS run performed at nominal 11 °C. On top of the scheme, grey bars indicate timespans where DRS valves were open (used during DRS runs, stand-by during LTP runs). Red markers at the bottom indicate noise-only long-lasting runs used for spectral analyses. Labels mark relevant LPF operations.

preliminary adjustments, and alignments. On DAL 62.4, the two vacuum enclosures containing the GRSs activated the CVMs, venting to space and unlocking the TMs. The area outside the CVM valve already had an open path to space, but the inner area was sealed. On the one side, venting allowed the system to start pumping the residual gas and lowering the pressure. Additionally, it unlocked the TMs, which were caged for launch. The GPRM allowed keeping them mechanically centered in the EH, providing $\sim$ N forces to the TM central recesses. On DAL 73, the GPRM unlocked, releasing the TMs. The GPRM was designed to release the TMs with small residual impulses, of the order of $\mu m\,s^{-1}$, which could be captured by the electrostatic actuation system. After an initial stabilization of the TMs inside their EHs, the drag-free mode was enabled. On March 1st, 2016, on DAL 88, science operations started.

Operations included many kinds of experiments. Science measurements consisted of pure noise-only runs, lasting between some days and a couple of weeks. During science runs, the purpose was the measurement and control of differential acceleration, by applying the force commanded by control loops without additional spurious stimuli, with stable experimental conditions. Other experiments and operations included system identification and calibration runs, station-keeping maneuvers, interferometer alignments, TM discharge, charge measurements, etc. In Figure 2.15, colored timespans represent the runs used in this work, as described in the following paragraphs. Runs are also listed in App. A, in which we give more details on the experimental conditions.

Control schemes. Various control schemes were applied during the LPF mission. During the first part of the mission, from March 2016 to June 2016, LTP ordinary operations took place, using the LTP drag-free control scheme, and cold gas thrusters. In June 2016, control was handed to NASA's ST-7 operations, which used the DRS control scheme, operated with the colloidal spacecraft thrusters, with the `DURLA` actuation scheme. In the figure, we have marked in red those runs which

operated in DRS conditions, and that we have used in this work. However, colloidal thrusters have been in a stand-by state for a longer time, even during some LTP runs. The state of the DRS valves is shown in the upper part of the figure.
In June 2016, the mission extension began, alternating LTP and DRS operations. DRS operations in 2017 (run Apr17A), however, required some cold gas thrusters to be activated, because of a problem with a colloidal thruster cluster.
On April 2017, LPF was de-orbited from L1. Operations went on till July 2017, when communications were definitively interrupted.

Actuation settings. Within the LTP operations, various actuation authority schemes were implemented, aiming at lowering the actuation noise. Some special "free-fall" experiments were also performed, during which actuation on TM2 was performed with impulses instead of continuous forces: however, we do not deal with those experiments in this work. Details about the authority schemes are given in Table A.3.
At the beginning of the operations, nominal (`NOM`) actuation was implemented, allowing for a maximum thrust on TM2 of 2.2 nN. However, gravitational compensation was so effective that such a large maximum force was not necessary. Therefore, actuation authority was progressively reduced to `RED` (600 pN TM2 authority), `RLA` (200 pN TM2 authority), `URLA` (50 pN TM2 authority), and , `UURLA` (50 pN TM2 authority, and reduced angular authority). The `UURLA` settings were applied from April 2016.

Thermal settings. LPF operated nominally at room temperature, nominally 22-24 °C at TM locations, for the majority of the mission (see App. A for the average temperature during runs). However, some variations were applied during the mission extension. Such variations led to important experimental results, from the point of view of low-frequency noise and transient events.
In January 2017, temperature at TM location was lowered to 11 °C, aiming at lowering the white noise plateau due to Brownian molecular motion. This led to one of the quietest long-lasting runs of the LPF mission, run Feb17B, which was used for analyses in [2], and which we often use as a reference.
After the de-orbiting maneuver, temperature was further lowered in April 2017, switching off all the heaters, aiming at the lowest temperature possible, ~ 0 °C. Four runs were performed at this temperature. At the end of May 2017, heaters were switched on again, rising temperature to 11 °C as first step, and 22 °C in June 2017.

Analyzed runs. In the following analyses, we use the set of runs listed in Table A.1. We have two main analysis directions: the spectral analysis of low-frequency force noise, and the time-domain analysis of spurious transients. Different subsets of runs are used for two purposes:

1. For the analysis of transient "glitches", we use the entire set in Table A.1. Some of the runs were performed with non-optimal actuation authorities, resulting in an increased noise level: however, the sensitivity was sufficient for the detection of transients. In that context, we will refer to *ordinary runs* (OR) as the ones performed at 22 °C or 11 °C, and to *cold runs* (CR) as the ones performed in May 2017, at about 0 °C. All of these are noise-only runs, except for a couple them, during which a fN sinusoidal stimulus was applied to TMs, without

interfering with our purposes of glitch identification.

2. For the analysis of low-frequency noise, we use a subset of 13 of those runs, and an additional DRS run, as marked in orange in Table A.1. Conditions of the LTP runs are detailed in Table A.2. In our analyses, run Nov16B is split into two segments, because of a micrometeoroid event occurring about two days before the end of the run. These (LTP) runs were performed with the `URLA/UURLA` actuation settings, and no particular external stimuli were applied during measurements. Runs in this subset last from 2.75 d (Jul16C) to 18.5 d (Dec16B). Long durations allowed to analyze frequencies down to 18 µHz, with 13 periodograms to be averaged also at the lowest frequency in the longest run[6].

Cold runs. Cold runs require some more details and explanations, as their role in the following analyses is relevant. Cooldown began on April 29th, 2017, bringing temperature down from 11 °C to the lowest possible, switching off all the heaters. The reason for a cooldown was simple: Brownian noise is proportional to temperature, and a first cooldown had already succeeded in improving the noise performance. Such a low temperature was however out of the mission design and specifications: thermometers on the EH saturated when reaching about 8 °C, hence being unusable below that temperature. Thermometers on the outer LTP bay worked well, measuring temperatures down to −1 °C in May 2017. In [44], it is estimated that temperatures at TMs locations went down to ~ 2 °C. Four runs were performed at this temperature (May17A to May17D), amounting to a total of 12.4 d.
During cold runs, differential acceleration measurements showed the occurrence of an increased number of transient "glitches", as described in Chapter 7. The high number of glitches, contributing both to the low-frequency noise and to the white level, spoiled noise measurements, resulting in a highly non-stationary behavior. Moreover, between run May17B and run May17C, a more serious event occurred, requiring the spacecraft to re-grab the TMs in *safe mode*. That event is presumably of mechanical origin [45]. After re-heating, the structure went on showing some non-thermal mechanical relaxation. More details later on, see Sec. 4.3.3 on page 88.

2.2 LISA Pathfinder: data processing, data series

Let's now proceed with the description of the main science measurements and the time series definitions. We give more details about the system dynamics, especially concerning the science differential measurements, and describe the post-processing pipelines used for data series production.

Raw data from instruments have been onboard low-passed, subsampled, and telemetered with a sampling frequency $f_s = 10$ Hz, hence allowing measurements up to the Nyquist frequency of 5 Hz. All the raw time series have been manipulated with LTPDA (LTP Data Analysis toolbox), the MATLAB toolbox developed for LPF data

[6]The choice of 18 µHz allows to have 100 µHz as fourth frequency, which is the lower bound of the LISA frequency band. Measurements at that frequency are taken with periodograms lasting 222222.2 s, following the scheme in Sec. 3.5.

analysis [46]. The first elaboration allows to re-cast all data interpolating them to the same time grid, allowing for further calculations and analyses.

In this section, we make use of spectral analysis tools. For now, we do not need sophisticated instruments, as here we only use Power and Amplitude Spectral Densities, PSDs, and ASDs. More details on the computing algorithms, definitions, and properties are given in Chapter 3. Whenever a PSD is represented as points with error bars, it has been computed at minimally-correlated frequencies with the estimation tools provided in that chapter; if it is a continuous line, it is instead computed with the standard Welch method.

2.2.1 Science measurements and system dynamics

The main LPF science measurement is the out-of-loop differential acceleration $\Delta g(t)$, i.e. the differential acceleration that the masses would show if no actuation loop were involved. The presence of the actuation forces must be taken into account in building Δg, as the loops play a relevant role in the dynamics of the spacecraft, TM1, and TM2, especially at sub mHz frequencies.
To understand the working principle, we go over the Newton equation behind it, restricting it along the x coordinate.

Assume that $\xi_1(t)$, $\xi_2(t)$, and $\xi_{\rm SC}(t)$ represent the coordinates with respect to an inertial reference frame of, respectively, TM1, TM2, and the spacecraft. These coordinates are related to the measured $x_1(t)$ and $\Delta x_{12}(t)$, through $x_1(t) = \xi_1(t) - \xi_{\rm SC}(t)$, and $x_{12}(t) = \xi_2(t) - \xi_1(t)$. The Newton equation along x, dropping the explicit dependence on time, reads:

$$\begin{cases} m\ddot{\xi}_1 + m\omega_1^2\left(\xi_1 - \xi_{\rm SC}\right) = mg_{e,1} + mg_{c,1} \\ m\ddot{\xi}_2 + m\omega_2^2\left(\xi_2 - \xi_{\rm SC}\right) = mg_{e,2} + mg_{c,2} \\ m_{\rm SC}\ddot{\xi}_{\rm SC} - m\omega_1^2\left(\xi_1 - \xi_{\rm SC}\right) - m\omega_2^2\left(\xi_2 - \xi_{\rm SC}\right) = \\ \qquad = m_{\rm SC}g_{e,{\rm SC}} + m_{\rm SC}g_{c,{\rm SC}} - mg_{c,1} - mg_{c,2} \end{cases} \tag{2.4}$$

Here, m and $m_{\rm SC}$ are the TM mass and the spacecraft mass, assuming TM1 and TM2 have equal masses. The coefficients ω_i^2 represent the gradient forces, expressed as stiffness terms. The first and the second equation link the left-hand side, representing the TM motion, with the right-hand side, containing the real external forces per unit mass (g_e), and the applied electrostatic control forces per unit mass (g_c). The third equation rules the spacecraft motion, dominated on the right-hand side by the external forces $m_{\rm SC}g_{e,{\rm SC}}$, and the control thrusts $m_{\rm SC}g_{c,{\rm SC}}$. These equations do not explicitly contain the control transfer functions, as those are implicitly included in the applied control forces.
Expanding the third equation with respect to the interferometer measurement x_1, one gets the spacecraft equation of motion, ruling the spacecraft motion with respect to the x_1 measurement, which is not our focus now.
Instead, from the first and second equations, one gets the equation ruling the differential displacement:

$$m\left(\ddot{x}_2 - \ddot{x}_1\right) + m\omega_2^2 x_2 - m\omega_1^2 x_1 = mg_e + mg_c \tag{2.5}$$

where g_e and g_c simply condense the differential forces, $g_e = g_{e,2} - g_{e,1}$ and $g_c = g_{c,2} - g_{c,1}$. Re-introducing the explicit dependence on time, this equation can be re-arranged, and represents the main LPF output,

$$\Delta g(t) = \Delta\ddot{x}_{12}(t) - g_c(t) + \Delta\omega_{12}^2 x_1(t) + \omega_2^2 x_{12}(t) \tag{2.6}$$

Summarizing, in Eq. (2.6):

- $g_c(t)$ represents the applied differential actuation force per unit mass along x, i.e. $g_c = g_{c,2} - g_{c,1}$. Following the control scheme described in Sec. 2.1.3, the drag-free mode should apply no force along the sensitive axis to TM1. Thus, $g_{c,1}$ is nominally 0. Nevertheless, the application of digitized actuation waveforms was affected by a rounding error which was not corrected in flight [47]. Hence, the real applied forces were computed in post-processing, also resulting in a small, though negligible, force per unit mass $g_{c,1}$. System identification experiments also identified amplitude calibration factors and delays in force applications ($\lambda_{1,2}$, $C_{1,2}$, as described in [24], which we here include in g_c for simplicity).

- Gradient forces, resulting from both gravitational field gradients and electrostatic actuation gradients, are modeled as elastic-like forces with negative stiffnesses. Such modelization and linearization are possible, as the system is held in an equilibrium state with the feedback force loops. Each TMi, within its EH, is subject to a force per unit mass $-\omega_i^2 x_i(t)$, where $x_i(t)$ is the displacement from the nominal position. Each stiffness ω_i^2 is contributed by a constant term, mainly gravitational, common to all runs, and an electrostatic term, proportional to the actuation force and torque authorities. Stiffnesses have been measured in system identification experiments [24]. With the `UURLA` settings, $\omega_1^2 = (-4.41 \pm 0.07) \times 10^{-7}\,\mathrm{s}^{-2}$ and $\omega_2^2 = (-4.62 \pm 0.04) \times 10^{-7}\,\mathrm{s}^{-2}$; these stiffnesses are contributed $\sim 10\%$ by electrostatic terms. Since the measurement is differential and based on x_1 and Δx_{12}, these stiffnesses combine as in Eq. (2.6), where $\Delta\omega_{12}^2 = \omega_2^2 - \omega_1^2$. Contribution from the $\Delta\omega_{12}^2$ term is negligible with respect to the ω_2^2 one, and, in `UURLA` runs, the latter is negligible with respect to the measured noise, as noticeable in Fig. 2.17.

- $\Delta x_{12}(t)$ is the time series of the x12 interferometer output (also indicated as $\Delta x_{12,\mathrm{OMS}}(t)$), translated from laser phase to differential TM motion. It is an in-loop quantity, which is used by the suspension loop as an error signal to the reference set point. Therefore, it is heavily influenced by the suspension loop transfer function, which acts mostly at low (sub-mHz) frequencies. At those frequencies, the PSD of $\Delta x_{12}(t)$ goes quickly to zero.

- Common-mode between the x1 and x12 was effectively rejected by the drag-free scheme, hence $\delta_{\mathrm{ifo},1}\ddot{x}_1(t)$ represents a negligible correction.

Figure 2.16 provides schematics of the suspension loop along x, with stiffness terms, as depicted in Eq. (2.6). The figure also shows the role of the GRS readout, which will be clarified later, and the role of the interferometer noise. Not shown in the picture is the role of the actuation noise, which is treated separately, among the other

noise sources. In Figure 2.17, we show as an example the PSD of the various terms

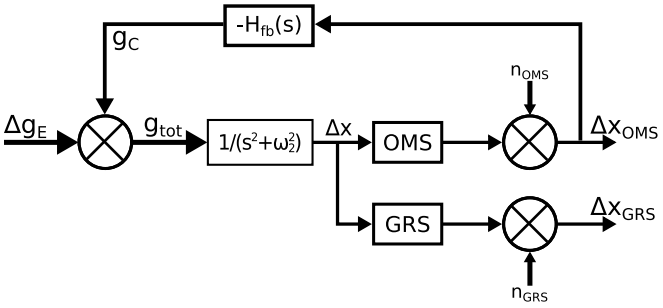


Figure 2.16: Schematics of the suspension loop along x, actuating forces on TM2 with respect to its relative position to TM1. Δg_E, the external forces acting on the TMs, translate into TM motion through the transfer function $1/(s^2+\omega_2^2)$, where s is the Laplace frequency. The interferometric readout $\Delta x_{\rm OMS}$ is fed back to the DFACS, which applies control forces per unit mass g_c, with a transfer function $-H_{\rm fb}(s)$. The GRS plays no role in the loop, hence providing a useful external independent readout of the TM position.

appearing in $\Delta g(t)$. Some of them have already been defined. Others are described in the following paragraphs. We note that LPF time series have been affected by transient events "glitches", treated in detail in Chapter 7. In the figure, we have deglitched the time series, i.e. we have fitted and removed the spurious transients that occurred during measurements, for clarity [3].

2.2.2 Correction of interferometer cross-talk

Data show a well-visible "bump" between 10 mHz and 0.1 Hz, evident in Fig. 2.17. That bump is not ascribable to the readout error of the interferometric measurement, nor to real forces acting on the TMs along the x axis. As explained in Sec. 2.1.2, that feature is due to pickup of acceleration along other degrees of freedom, namely $\ddot{\phi}(t)$, $\ddot{\eta}(t)$, $\ddot{y}(t)$, $\ddot{z}(t)$, coupling to the sensitive axis because of misalignment. Post-processing decorrelation has proved effective in reducing that spurious coupling [24, 48], as noticeable in Fig. 2.17, in the lower panel. More recently, steps forward are underway in modelization [49].

2.2.3 Correction of inertial effects

Time series resulting from Eq. (2.6) also show a "bump" below 0.5 mHz, which suggested correlation to some sort of low-frequency signal. It has been shown [1–3] that such a feature was caused by the spacecraft platform being non-inertial, hence coupling TM acceleration to the spacecraft angular jitter.

The ultimate reason has to be found in the attitude control system. The DFACS controlled the spacecraft attitude with autonomous star trackers, maintaining a low drift to ensure correct antenna pointing. Nevertheless, the star tracker readout was noisy at low frequencies, causing the spacecraft to jitter, generating a jittering angular velocity and a jittering angular acceleration. As previously done in LPF analyses, we

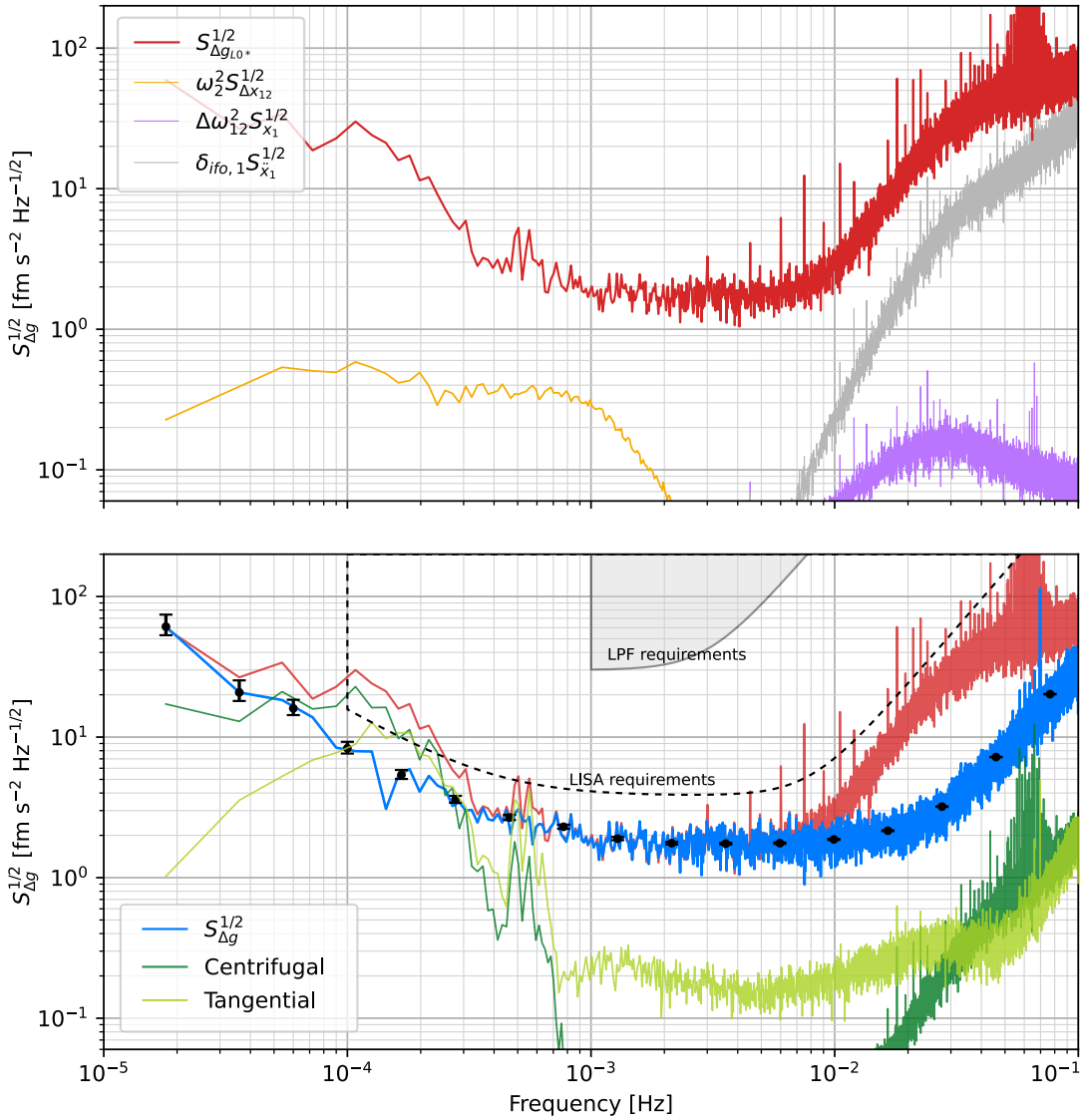


Figure 2.17: LPF on-ground data processing, example with run Feb17B, one of the quietest run performed. The upper panel shows the first step of data production, resulting from the interferometric measurement, feedback forces, and force gradients. The red line represents the ASD of $\Delta g_{\mathrm{L0}^*}(t)$, namely $\Delta g(t)$ as computed as in Eq. (2.6), and deglitched. The yellow, purple, and grey curves represent the contributions from force gradients and the pickup from the x1 interferometer.
In the lower panel, we show the ASD of $\Delta g_{\mathrm{L0}^*}(t)$, and $\Delta g(t)$ after the subtraction of inertial effects and spacecraft movement pickup. We also show the single contributions from centrifugal and tangential forces, and the LISA/LPF requirements.

subtract such effects from Δg. Neglecting this source in the LPF residual noise results is an appropriate operation, since this noise is not expected in LISA data: angular control will be based on the differential waveform sensing between the local and the distant lasers, making angular control much more stable [12].

Inertial forces contribute in a twofold manner:

1. A noisy centrifugal force, proportional to the jittering angular velocity $\boldsymbol{\Omega}(t)$. Such a force is purely differential, pushing both the TMs outwards. It can be expressed as

$$\Delta g_{\Omega} = \left(\boldsymbol{\Omega}(t) \times \boldsymbol{\Omega}(t) \times \boldsymbol{r}_{12}\right)_x = -(\Omega_{\phi}^2(t) + \Omega_{\eta}^2(t))\, r_{12}, \tag{2.7}$$

   where $\boldsymbol{\Omega}(t)$ is the spacecraft angular velocity, and $\boldsymbol{r}_{12} = (r_{12}, 0, 0)$ the vector joining the two TM centers, nominally 0.376 m along the x direction. Spacecraft angular velocity is deduced from the star tracker in-loop measurement and the applied torques. More precisely, the angular velocity $\Omega(t)$ is made of a quasi-DC component $\Omega_{\mathrm{DC}}(t)$, measured by the star trackers, and an in-band component $\Omega_{\mathrm{B}}(t)$, which we compute integrating in time the applied common-mode torques, along ϕ and η. The result is quadratic, as in Eq. (2.7). Hence, the quasi-DC component directly enters into the measurement frequency band. Overall, Eq. (2.7) can effectively be approximated as:

$$\Delta g_{\Omega} \sim -2\Omega_{\phi,\mathrm{DC}}(t)\Omega_{\phi,\mathrm{B}}(t)\, r_{12} - 2\Omega_{\eta,\mathrm{DC}}(t)\Omega_{\eta,\mathrm{B}}(t))\, r_{12} \tag{2.8}$$

   This effect is calculated and subtracted, as shown in Fig. 2.17.

2. A noisy tangential "Euler" force, proportional to the jittering angular acceleration $\dot{\boldsymbol{\Omega}}(t)$, expressed as

$$\Delta g_{\dot{\Omega}} = \left(\dot{\boldsymbol{\Omega}}(t) \times \boldsymbol{r}_{12}\right)_x = -\dot{\Omega}_{\phi}(t)\delta y + \dot{\Omega}_{\eta}(t)\delta z \tag{2.9}$$

   This effect is not quadratic, hence the quasi-DC angular velocity does not have any influence on it. If the TMs centers were perfectly aligned, such a term would not contribute, as $\delta y = \delta z = 0$, but it is not the real case. Euler tangential acceleration was subtracted with a frequency-domain fit to the model in Eq. (2.9), extracting the angular misalignments $\delta\phi = \delta y / r_{12}$, and $\delta\eta = \delta z / r_{12}$ [3, 24]. A global fit model was used, dividing the LPF mission into epochs, depending on the TM alignment configurations. Angular mismatches were found to be of the order of 10-100 µrad, depending on the TM alignment epoch. After estimation of the misalignment angles, this effect was computed and subtracted, as shown in Fig. 2.17, lower panel. Inertial forces contributed to more than 90% of the total measured noise at 0.1 mHz.

2.2.4 Deglitching

Transient acceleration events will be a matter of discussion in Chapter 7. It is however necessary to introduce them here, as deglitching is an important operation that

also affects spectral analysis. After subtracting all the effects listed above, transient acceleration glitches were identified in experimental runs. Such glitches could be effectively modeled, fitted, and subtracted from data [2, 3]. All the time series presented in Fig. 2.17 have been deglitched.

2.2.5 Angular differential acceleration time series

Analogously to the linear out-of-loop differential acceleration Δg, we have built the angular out-of-loop differential acceleration time series along the η and ϕ directions, $\Delta\gamma_\eta(t)$ and $\Delta\gamma_\phi(t)$. Angular displacement was subject to a different control scheme than the linear one [40], using suspension loops on both TMs and actuating torques based on interferometric readouts. Pitch and yaw TM motion (η and ϕ) were both controlled interferometrically, hence the out-of-loop angular acceleration can be built in first approximation by combining the interferometric measurements and the known applied torques. Potentially, high-precision measurements would be available for both test masses, independently. However, angular motion is dominated by the spacecraft's angular jitter, which shows up as a high-intensity common-mode acceleration. Therefore, it is quite impractical to retrieve the angular acceleration of each TM. We analyze the differential angular acceleration, much quieter than the single ones.

To get pure differential motion, we also decorrelate the residual common-mode motion [3]. Finally, the angular acceleration time series (torque per unit moment of inertia) are defined as:

$$\Delta\gamma_\phi = \ddot{\phi}_2 - \ddot{\phi}_1 - (N_{\phi_2} - N_{\phi_1})/I_{zz} - (\omega^2_{\phi_1}\phi_1 - \omega^2_{\phi_2}\phi_2) - \alpha_{\overline{\gamma}_\phi}\overline{\gamma}_\phi, \tag{2.10}$$

where N_{ϕ_i} is the torque applied by the x-ϕ electrodes to TMi, I_{zz} is the moment of inertia about the z axis, $\omega^2_{\phi_i}$ the angular stiffnesses (whose terms play a negligible role). Similar equations hold for η.

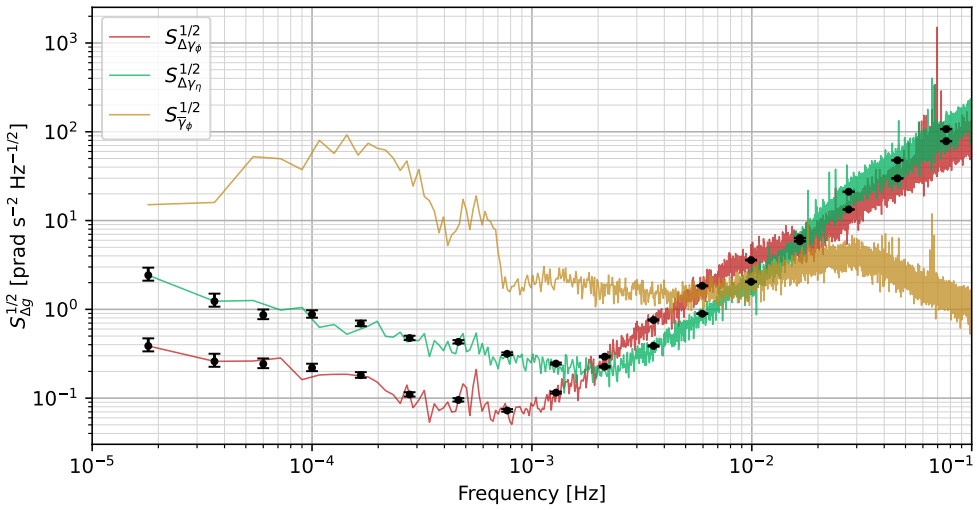


Figure 2.18: Example of the power spectra of the differential angular accelerations $\Delta\gamma_\phi$ and $\Delta\gamma_\eta$, during the noise run Feb17B. For comparison, we also show the common-mode $\overline{\gamma}_\phi$, along ϕ, which has also been decorrelated from the $\Delta\gamma_\phi$ time series.

Figure 2.18 shows examples of the power spectra of the differential angular accelerations $\Delta\gamma_\phi$ and $\Delta\gamma_\eta$, during the Feb17B run. It also shows the power spectrum of the common-mode acceleration $\overline{\gamma}_\phi$, whose power is $\sim 10^4$ times higher.

We will note in Chapter 7 that some glitches showed a small torque. For those events, we perform the deglitching operation also for the $\Delta\gamma$ time series.

2.2.6 GRS displacement timeseries

The position of TM1 is driven by the interferometer x1, which holds it centered with respect to the optical bench position. The same holds for TM2, whose position is centered with respect to TM1, in the x, η, and ϕ coordinates. Therefore, the position of the TMs is not constrained with respect to their EH, and a movement of the EH itself would not induce a feedback force on the TM. As schematized in Fig. 2.19, a deformation in the structure holding the two GRS would not be detected by the interferometer loop. Assume that EH1 and EH2 move from the nominal position with respect to the OB, respectively by the amounts $x_{1,\mathrm{GRS}}$ and $x_{2,\mathrm{GRS}}$; the interferometric measurements, $x_{1,\mathrm{OMS}}$ and $\Delta x_{12,\mathrm{OMS}}$, would not sense the displacement.

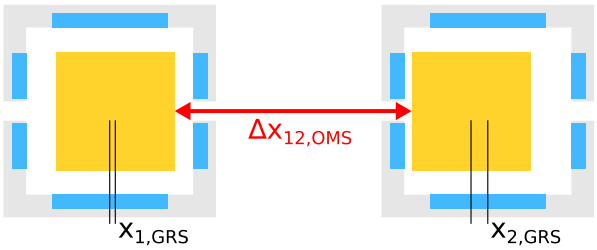


Figure 2.19: Definition of Δx_{GRS}, as in Eq. (2.11). The relative positions TM1-to-S/C and TM2-to-TM1 are controlled by the interferometers, rigidly bound to the OB. The loop keeps the TMs in their nominal position with respect to the OB. Hence, the GRS provides capacitive sensing of the displacement of the GRS structure with respect to the TMs, which is not available with interferometric measurements. TM size is not to scale, and EH motion is exaggerated, for visualization purposes.

As noted in Fig. 2.16, the GRS capacitive sensors are not involved in the suspension loops along these degrees of freedom, and are intrinsically rigidly moving with the EHs. Therefore, those can be used as measurements of the instrument distortion. We build the time series Δx_{GRS}, defined by

$$\Delta x_{\mathrm{GRS}} = x_{2,\mathrm{GRS}} - x_{1,\mathrm{GRS}}, \tag{2.11}$$

where $x_{1,\mathrm{GRS}}$ and $x_{2,\mathrm{GRS}}$ are, respectively, the positions of TM1 and TM2 within their EHs. The intrinsic noise level of this time series is generally higher than that affecting $\Delta x_{12,\mathrm{OMS}}$, simply because the capacitive sensing is much noisier than the interferometric one. At low frequencies, where the TM motion dominates, the PSD of the two sensing methods is comparable.

Distortion stiffness. We note that the mechanical distortion pattern along the x axis would however leak into the differential motion. The mechanism is similar to the one coupling the x1 and x12 readouts to Δg, caused by force gradients. The

GRS, indeed, generates gravitational and electrostatic force gradients on the TMs, generating elastic-like forces as the EHs move. We model the distortion-induced acceleration $\Delta g_{\rm dist}$ as

$$\Delta g_{\rm dist} = -\omega_d^2(\Delta x_{\rm GRS} - \Delta x_{\rm OMS}), \tag{2.12}$$

where ω_d^2 is a stiffness coefficient. From $\Delta x_{\rm GRS}$, we have removed $\Delta x_{\rm OMS}$ to remove true TM motion. However, ω_d^2 is different from ω_1^2 and ω_2^2: the two cases are not equivalent, mainly because in the latter the OB does not move with respect to the TMs.

We evaluate the stiffness following what done in [24], using the coefficients computed in [50]. Let's assume that X_i and x_j are respectively the off-nominal position of GRSi and TMj with respect to the spacecraft. $\omega_{i,j}^2$ is the stiffness caused by GRSj on TMi, and $\omega_{{\rm SC},j}^2$ the stiffness caused by the rest of the spacecraft on TMi. We notice that the stiffnesses ω_i^2 in Eq. (2.6) are given by $\omega_i^2 = \omega_{i,1}^2 + \omega_{i,2}^2 + \omega_{{\rm SC},i}^2$. The gradient forces are given by

$$\begin{aligned}\Delta g_{\rm grad} =& -\omega_{2,2}^2\,(x_2 - X_2) - \omega_{{\rm SC},2}^2 x_2 - \omega_{2,1}^2\,(x_2 - X_1) + \\ &+ \omega_{1,1}^2\,(x_1 - X_1) + \omega_{{\rm SC},1}^2 x_1 + \omega_{1,2}^2\,(x_1 - X_2) = \\ =& -(\omega_2^2 x_2 - \omega_1^2 x_1) + (\omega_{2,2}^2 - \omega_{1,2}^2)X_2 - (\omega_{1,1}^2 - \omega_{2,1}^2)X_1\end{aligned} \tag{2.13}$$

The last two terms are the contribution of the EH displacement, $\Delta g_{\rm dist}$. We have already taken into account the first term, which is the contribution of TM displacement. Assuming that $(\omega_{2,2}^2 - \omega_{1,2}^2) \sim (\omega_{1,1}^2 - \omega_{2,1}^2)$, we conclude that $\omega_d^2 \sim (\omega_{2,2}^2 - \omega_{1,2}^2)$. Taking into account both gravitational and electrostatic terms, we estimate the effective stiffness $\omega_d^2 = (-3.32 \pm 0.05) \times 10^{-7}\,{\rm s}^{-2}$.

2.2.7 Considerations on measurements

It is well noticeable in Figure 2.17 that noise is mainly made of three branches: a rightmost part, scaling with frequency as $S_{\Delta g}(f) \propto f^4$, a leftmost part, scaling as $S_{\Delta g}(f) \propto f^{-2}$, and a central part, frequency-independent.

1. The right branch, scaling as $S_{\Delta g}(f) \propto f^4$, is confidently ascribed to interferometer position readout noise [13, 33], and it is not of interest to our purposes.

2. The central branch, frequency-independent between $\sim 1\,$mHz and $\sim 10\,$mHz, is confidently ascribed to Brownian noise and outgassing. We describe it in Chapter 5.

3. The left branch, known to be scaling as $S_{\Delta g}(f) \propto f^{-2}$ [2, 3], is instead affected by a number of sources and mechanisms. In Chapters 4, 6, and in the discussion 8, we provide analyses of acceleration noise in the sub-mHz band, investigating the nature of the observed noise.

The next Chapter 3, instead, runs on a different track. It deals with spectral estimation methods, cross-correlation analysis, and frequency-domain decorrelation tools, which we use extensively in the following analyses.

3 Spectral estimation and decorrelation

Multivariate spectral estimation,
time series coherence and signal decorrelation

This chapter runs on a different track from the previous one and the following ones. We break the discussion on the LPF noise, to describe the methods we used – and developed – for the spectral analysis of time series.
Note: this chapter contains many equations; for clarity, we mark with an orange box the most relevant ones.

There are multiple reasons behind the necessity of developing statistical tools for the estimation of spectral quantities. First of all, the necessity of analyzing very low frequencies, and subsequently dealing with just a few periodograms for spectral analyses. The inability to average over a large number of periodograms makes the spectral quantities very far from being Gaussian-distributed, requiring more precision when giving uncertainties. Second, the necessity of analyzing the uncertainties of spectral quantities, both commonly used and non-standard, in looking for the fine details of coherences among time series. Detailed knowledge of power spectra statistics is also required for analyzing the stationarity and the long-term instrument stability. We also propose some practical applications of the studied statistics, and a decorrelation method to identify the contribution of supplementary time series to a main measurement. These tools are explicitly developed for LPF analyses, but can be easily generalized.

This chapter follows and expands the analysis methods presented in [2, 3, 51], extending the spectral analyses to multi-channel measurements. The methods here presented have been implemented in the analyses that will be presented in [5], also discussed in Chapter 6.

In Sec. 3.1 we introduce the basic definitions of spectral analysis, to establish the grounds for the following elements. In Sec. 3.2 we present the complex Wishart distribution and describe the statistics of (cross-)power spectra, coherences, and multiple coherence. In Sec. 3.4 we present our Bayesian single-frequency decorrelation method, along with the model definition, and the posterior parameter distributions. Finally, in Sec. 3.6, we extend the single-frequency method to a larger set of frequencies, and provide numerical tests validating the presented methods. The choice of quasi-independent frequencies is described and justified in Sec. 3.5. Additionally, in Appendix B we illustrate an algorithm for the generation of random complex-Wishart-distributed (cross-)power spectral matrices, and provide some more details.

3.1 Spectral analysis and basic definitions

Let's assume that the measurement consists in p real-valued time series, which have been simultaneously and synchronously measured with a fixed sampling time T. These time series, which we call $x_i(t)$, represent measurements of different experimental quantities. As in the LPF case, let's additionally assume that the sampling frequency $f_s = 1/T$ is sufficiently high to neglect any aliasing of higher-frequency components into the low-frequency band. Generally, we assume that such p time series are real-valued stochastic processes, just because this is the case for experimental data. However, data could as well be complex-valued without major modifications to the following analysis.

If the measured time series are stochastic processes, stationary within the measurement time and Gaussian, then they can be characterized by a single quantity defining the spectral properties, known as the Cross Power Spectral Density (CPSD) matrix, $\mathbf{\Sigma}(f)$. The CPSD matrix is commonly defined as the Fourier transform of the autocorrelation matrix, as described in [52]. At a given frequency f, the CPSD matrix is built such that each diagonal element $S_{ii}(f)$ is the Power Spectral Density (PSD) of the i-th stochastic process, and the off-diagonal elements $S_{ij}(f)$ are the cross-PSDs among the i-th and j-th processes.

The theoretical CPSD matrix is characterized by some simple properties,

1. The CPSD matrix is Hermitian, by definition. Therefore, its diagonal elements, the PSDs, are real-valued. Moreover, by definition, those elements are not negative.

2. The CPSD matrix is positive definite. Indeed, it exists a coordinate change that reduces the cross-correlations to zero, diagonalizing the matrix. The diagonal elements of a such-defined matrix would be PSDs on their turn, and hence real- and positive-valued.

3. The diagonal elements of the standardized CPSD matrix are known as the cross-coherences between the p time series, defined as

$$\rho_{ij}(f) = \frac{S_{ij}(f)}{\sqrt{S_{ii}(f)\, S_{jj}(f)}} \tag{3.1}$$

   The square modulus of this quantity, $|\rho_{ij}|^2$, is often indicated as Magnitude Square Coherence (MSC).

4. The *multiple* coherence R^2 associated with the CPSD matrix $\mathbf{\Sigma}$, with respect to the first time series, is defined as

$$R^2 = 1 - (\Sigma_{11}\Psi_{11})^{-1}\,, \tag{3.2}$$

   where $\mathbf{\Psi} = \mathbf{\Sigma}^{-1}$. We will prove that the multiple coherence coincides with the total contribution fraction to noise of the first time series by the other ones.

3.1.1 Discrete time series and CPSD estimators

Given the discrete-time nature of the time series, we write $x_i[n] = x(nT)$, meaning that $x_i[n]$ is the n-th sample of the i-th time series[1], measured at time $t = nT$.

The building block at the basis of the whole spectral theory is the Discrete Fourier Transform (DFT). For our purposes of spectral estimation, we weight data with a spectral tapering window $w[n]$, such that $\sum_n (w[n])^2 = 1$. As will be clarified later, the ultimate purpose of windowing is to prevent spectral leakage among adjacent frequency samples, preventing the onset of numerical artifacts that would otherwise be generated. The modified DFT, performed on an N-samples long data stretch, is defined as

$$X_i[k]^{(N)} \equiv \sqrt{\frac{T}{N}} \sum_{n=0}^{N-1} x_i[n] w[n] e^{-2\pi i n k/N}. \tag{3.3}$$

Here k is the spectral index, spanning between 0 and N. The spectral index is related to frequency, via $f = k/NT$. If data are real-valued, $X_i[k] = X_i[N-k]^*$, hence only $N/2$ samples are really independent. The maximum measurable frequency is known as the Nyquist frequency, and it is ultimately set by the sampling time, $f_{\mathrm{Nyq}} = 1/2T$.

3.1.2 (cross-)periodograms and power spectral density estimation

Another quantity that is strictly related to the modified DFT of the experimental data is the one-sided experimental cross-periodogram between the i-th and the j-th series, defined as

$$\hat{S}_{ij}^{(N)}[k] \equiv 2 X_i[k]\, X_j^*[k] \tag{3.4}$$

The Wiener-Khinchin theorem relates this quantity to the CPSDs and PSDs. It is indeed a well-known result in spectral theory that the cross-periodogram as in Eq. (3.4) is an unbiased estimator of the one-sided CPSD S_{ij}, provided that the spectral window is appropriately chosen and that the spectrum does not contain strong lines in the neighborhood of the frequency $f = k/NT$ (more details in Sec. 3.1.3). This is the case for the time series that we are considering at very-low frequencies, hence it is safe to use Eq. (3.4) as a basis for building a spectral estimator, as it is done in the Welch method.

3.1.3 The Welch PSD estimation method

The standard tool used for spectral estimation is the Welch method [53], which provides a way to reduce the variance of the previously-defined CPSD estimator, by averaging over many periodograms. It implicitly assumes that the stochastic process is stationary and its spectrum does not change during the whole measurement. The Welch method takes advantage from splitting a long time series into shorter stretches, defined so they overlap. Our algorithm, presented later, slightly modifies the concept of PSD evaluation at different frequencies, but preserves the estimation methods of the Welch algorithm. As standard choice in spectral analysis, we use 50%-overlapping stretches, which does not introduce a significant correlation between adjacent periodograms but

[1] We use square brackets to highlight the discreteness of the time series.

allows to increase the number of available periodograms, reducing the variance of the (C)PSD estimation. All in all, if M periodograms are available, the (C)PSD estimator at frequency $f = k/NT$ is

$$\hat{S}_{ij}(f = k/NT) = \frac{2}{M} \sum_{m}^{M} X_{i,(m)}[k]\, X^*_{j,(m)}[k] \tag{3.5}$$

It is clear that the spacing between adjacent frequencies – i.e. the frequency resolution and the lowest achieved frequency – is set by $1/NT$, since the stretch length is fixed to N samples. Hence, the lowest frequency is set by the total stretch duration $NT = \Delta T$. A severe limitation of the Welch method is that it is heavily dependent on the choice of the stretch length: on the one hand, increasing N allows to go down in the lowest achieved frequency and improve resolution; on the other hand, increasing N worsens the precision of the CPSD estimator, increasing its variance. A tradeoff is then required in the choice of N, as lowering the lowest frequency increases the error on the higher frequencies. Additionally, as noted in [51], adjacent frequencies show a certain level of correlation, because of the finite length of the data stretches. The lowest frequency that must be taken depends on the choice of the spectral window, varying between the second and the fifth one, for the common window choices. This allows to avoid spurious leakage of the DC ($f = 0$) term into the CPSD estimation.

We propose an alternative approach, proposing to shorten the stretch length for higher frequency analysis, to overcome the limitations imposed by the fixed choice of the stretch length in the Welch method.

Tapering windows, spectral leakage

As previously stated, the implementation of spectral tapering windows is essential to avoid leakage between adjacent frequency bins. Without windowing with an appropriately chosen function, the resulting periodograms would be affected by their finite length and abrupt truncation, having the same effect as a signal going abruptly to zero, which is not the reality.
It can be proved that, in case there is no aliasing from high frequencies (i.e., the sampling time is high enough that the cross-spectrum above the Nyquist frequency is negligible), the expectation value of the (C)PSD estimator $\hat{S}_{ij}^{(N)}$ would converge to

$$\left\langle \hat{S}_{ij} \right\rangle^{(N)} (f = k/NT) = \int_{-\infty}^{+\infty} S_{ij}(f) \left| W\left(k/NT - f\right) \right|^2 df, \tag{3.6}$$

that is the convolution between the theoretical one-sided (C)PSD S_{ij} and the Fourier transform of the spectral window $W(f)$. Hence the latter must be chosen so that it is as close as possible to a Dirac-δ distribution, to make the estimator converge to S_{ij}.

We do not go deeper into the discussion on the tapering windows, as it would be much wider, and already well-known. We implement the 4-term Blackman-Harris window [54], which allows for high side-lobe suppression ($-92\,\mathrm{dB}$ with respect to the main lobe) and minimal overlap correlation (3.8% for 50% overlapping segments). In Sec. 3.5, we also analyze the degree of correlation between nearby frequencies.

3.2 Periodogram statistics and probability distributions

Let's now discuss, at a single frequency f and hence at a single spectral index $k = fNT$, the properties of the probability distribution ruling the statistics of the experimental CPSD matrix, and the related quantities. In this Section, for clarity, we drop the explicit index k. Here we introduce the complex Wishart distribution, which is the key distribution in our spectral analysis.

3.2.1 The Complex Wishart distribution

Let's consider that the array $\boldsymbol{X}$, built with the modified DFTs according to Eq. (3.3), is a sample from a complex p-variate Gaussian random variable. This holds because the considered time series are Gaussian stochastic processes. A complex p-variate Gaussian random variable is simply a p-tuple complex variable, whose real and imaginary parts constitute a $2p$-variate Gaussian distribution. If the complex covariance is given by $\boldsymbol{\Sigma}$, then the probability distribution of $\boldsymbol{X}$ is given by

$$p(\boldsymbol{X}|\boldsymbol{\Sigma}) = \frac{1}{\pi^p |\boldsymbol{\Sigma}|} \exp\left(-\boldsymbol{X}^H \boldsymbol{\Sigma}^{-1} \boldsymbol{X}\right), \tag{3.7}$$

where $(\cdot)^H$ represents the conjugate transpose. The conditional dependence is explicitly written for clarity. It is clear that a set of M DFTs (over M independent stretches) would follow a joint probability given by the product of distributions like Eq. (3.7). From this, it is possible to derive a distribution for the CPSD matrix, as the one that we are about to describe.

Let's define another matrix, $\boldsymbol{W}$, as the sum of the CPSD matrices computed on M independent data stretches. In terms of $\boldsymbol{X}$, its elements would be defined by

$$W_{ij} = \sum_{m=1}^{M} X_{i,(m)} X^*_{j,(m)} \tag{3.8}$$

Goodman [55] showed that it is possible to derive the joint probability distribution of the elements of the matrix $\boldsymbol{W}$, that is known as the *Complex Wishart distribution*, as it extends the definition of the real Wishart distribution. The complex Wishart distribution rules the distribution of the matrix $\boldsymbol{W}$, built averaging over M periodograms of p time series.

$$p(\boldsymbol{W}|\boldsymbol{\Sigma}, M) = \frac{|\boldsymbol{W}|^{M-p}}{\widetilde{\Gamma}_p(M)|\boldsymbol{\Sigma}|^M} \operatorname{etr}\left[-\boldsymbol{\Sigma}^{-1}\boldsymbol{W}\right], \quad M \geq p, \tag{3.9}$$

where $|\cdot|$ represents the determinant, etr the exponential trace $\operatorname{etr}(\cdot) = \exp(\operatorname{tr}(\cdot))$, and $\widetilde{\Gamma}_p(M)$ is the multivariate complex Gamma function, defined by

$$\widetilde{\Gamma}_p(M) = \pi^{\frac{1}{2}p(p-1)} \prod_{i=1}^{p} \Gamma\left(M - i + 1\right).$$

We indicate that $\boldsymbol{W}$ follows the complex Wishart distribution as $\boldsymbol{W} \sim \mathcal{CW}(\boldsymbol{\Sigma}, M)$. In

the following, we elaborate on the complex Wishart distribution, which is at the basis of the PSD and MSC estimates, as well as the decorrelation method. In Appendix B.1, we also provide an algorithm to draw random samples of complex-Wishart-distributed matrices, which can be applied in noise simulations.

Note on applicability and degeneracy

The complex Wishart distribution, Eq. (3.9), is defined on the domain of the positive-definite matrices $\boldsymbol{W}$, with full rank p. This means that it is valid only if the number of averaged periodograms is greater than or equal to the number of time series, $M \geq p$. Otherwise, the rank of the matrix $\boldsymbol{W}$ would be M, and the condition of positive definiteness would decay, making $|\boldsymbol{W}| = 0$. Indeed, $\boldsymbol{W}$ would be positive semidefinite, with $p - M$ null eigenvalues. In the degenerate case, the complex Wishart distribution would reduce to the *singular* complex Wishart distribution [56], in the following equation.

$$p(\boldsymbol{W}|\boldsymbol{\Sigma}, M) = \frac{\pi^{M(M-p)}|\boldsymbol{\Lambda}|^{M-p}}{\widetilde{\Gamma}_p(M)|\boldsymbol{\Sigma}|^M} \operatorname{etr}\left[-\boldsymbol{\Sigma}^{-1}\boldsymbol{W}\right], \quad M < p, \tag{3.10}$$

which is valid on the domain of the $p \times p$ positive semidefinite Hermitian matrices of rank M. Here, $\boldsymbol{\Lambda}$ is the matrix of the non-zero eigenvalues of $\boldsymbol{W}$. We will employ this distribution in a particular case later on, for time series decorrelation.

3.2.2 Sample PSD distribution

It is straightforward to note that, in the one-dimensional case, the CPSD matrix reduces to the PSD of the considered time series. Then, the Wishart distribution can then be used to derive the PSD distribution of a single time series. We note that, generally speaking, the diagonal elements of the $\boldsymbol{W}$ matrix should individually follow the same statistics as the one-dimensional PSD. This is justified by the obvious observation that measurements can not affect each other, and the measurement of the entire set of p time series can not affect the single i-th time series. The statistics governing the PSD is already known in literature, as for example it has already been described in [51]: if the theoretical PSD is S, and Π the sample[2] PSD averaged over the M periodograms, then Eq. (3.9) reduces to the classical well-known result

$$p\left(\Pi|S\right) = \frac{M^M \Pi^{M-1}}{\Gamma(M) S^M}\, e^{-M\Pi/S} \sim \Gamma(M, M/S), \tag{3.11}$$

where $\Gamma(\cdot,\cdot)$ is the Gamma distribution, and its entries are the shape and scale parameters. This result can also be restated using the chi-square distribution: the random variable $y = 2M\Pi/S$ is indeed distributed as a chi-square with $2M$ degrees of freedom. The distribution in Eq. (3.11) is not biased, i.e. its mean value is the theoretical PSD S, and its variance obviously reduces by increasing the number of averaged periodograms, being proportional to $1/M$. In Fig. 3.1(left), we show some

[2]We refer to the *sample* or *experimental* PSD, to make a clear distinction to the *theoretical* PSD, that will be analyzed later.

sample PSD distributions, with the same theoretical PSD $S = 1$, but a varying number of periodograms. In Fig. 3.1(right), we represent the sample PSD distribution with $M = 10$, compared to samples generated from a random Wishart distribution, generated as described in Appendix B.1. We note that the distributions agree, according to a Kolmogorov-Smirnov test.

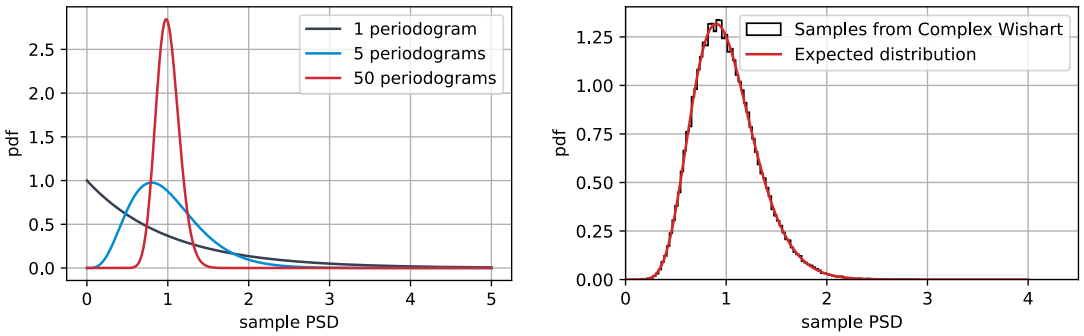


Figure 3.1: (left) Samples of the sample (experimental) PSD distributions, with theoretical PSD $S = 1$, with varying number of periodograms. We highlight that the distribution converges to the normal distribution for high numbers of averaging periodograms. (right) Sample PSD distribution with $M = 10$, compared to samples generated from a random Wishart distribution, generated as described in Appendix B.1. The distributions agree, according to a Kolmogorov-Smirnov test.

3.2.3 Sample cross-coherence distribution

Similar reasoning applies to the distribution of the sample MSCs, $|\hat{\rho}_{ij}|^2$. It can be proved that the probability distribution of this quantity has an analytical expression [55], which depends on the theoretical MSC $|\rho|^2$ and the number of averaged cross-periodograms, but does not depend on the PSDs. As reported in [57], the sample multiple coherence and the sample conditional coherence coincide in the bi-dimensional case, and even coincide with the magnitude-squared sample cross-coherence, $|\hat{\rho}|^2$. The probability distribution reads

$$p(|\hat{\rho}|^2||\rho|^2) = (M-1)(1-|\hat{\rho}|^2)^{M-2}(1-|\rho|^2)^M \times {}_2F_1(M, M, 1, |\hat{\rho}|^2\,|\rho|^2) \qquad (3.12)$$

where ${}_2F_1$ represents the Gauss' hypergeometric function. As per the PSD case, in Fig. 3.2 we show the sample MSC distribution, and a comparison to samples generated from a random Wishart distribution, generated as described in Appendix B.1, with theoretical MSC $|\rho|^2 = 0.8$, and number of cross-periodograms $M = 10$.

3.2.4 Multiple coherence distribution

The case of the multiple coherence is similar to that of the MSC. It can be proved [55] that the sample multiple coherence $\hat{R}^2$, defined as in Eq. (3.2), is described by an analytical distribution depending only on the theoretical R^2 and the number of averaged periodograms. If $\boldsymbol{W}$ is distributed as $\mathcal{CW}(\boldsymbol{\Sigma}, M)$, then the distribution of

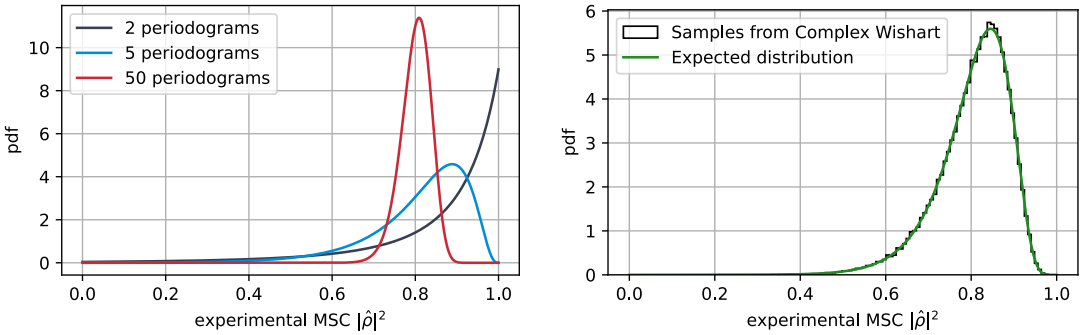


Figure 3.2: (left) Samples of the sample (experimental) MSC distributions, with theoretical MSC $\rho^2 = 0.8$, and varying number of periodograms. (right) Sample MSC distribution with $M = 10$, compared to samples generated from a random Wishart distribution, generated as described in Appendix B.1.

the sample multiple coherence $\hat{R}^2$ is

$$p(\hat{R}^2|R^2) = \frac{\Gamma(M)}{\Gamma(p-1)\Gamma(M-p+1)} (\hat{R}^2)^{p-2}(1-\hat{R}^2)^{M-p} \times \\ \times (1-R^2)^M \, {}_2F_1(M, M, p-1, \hat{R}^2 \, R^2) \tag{3.13}$$

Where R^2 is the theoretical multiple coherence, computed applying Eq. (3.2) to the matrix $\boldsymbol{\Sigma}$. We note that, in the 2-D case, the multiple coherence and the MSC coincide.

3.2.5 Schur complement and time series decorrelation

An interesting result comes from the block decomposition of the sample CPSD matrix. Let's assume that $\boldsymbol{W}$ is distributed as $\mathcal{CW}(\boldsymbol{\Sigma}, M)$, and let's decompose it into four blocks as follows:

$$\boldsymbol{W} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{B}^H & \boldsymbol{C} \end{pmatrix} \tag{3.14}$$

where $\boldsymbol{A}$ is a $q \times q$ matrix, $\boldsymbol{C}$ is $r \times r$ and $\boldsymbol{B}$ is $q \times r$. A known definition is the *Schur complement* of the block $\boldsymbol{C}$ in $\boldsymbol{W}$, and it is denoted as

$$\boldsymbol{W}/\boldsymbol{C} = \boldsymbol{A} - \boldsymbol{B}\boldsymbol{C}^{-1}\boldsymbol{B}^H \tag{3.15}$$

The interesting point is that, in case $\boldsymbol{A}$ is the PSD of a single time series – in the LPF case, the differential acceleration –, then the Schur complement of $\boldsymbol{C}$ in $\boldsymbol{W}$ coincides with the PSD of the residual noise after the decorrelation of the r time series in $\boldsymbol{C}$. This is the starting point of the single-frequency decorrelation method, which we discuss in more detail in Sec. 3.4.

If $\boldsymbol{W}$ is distributed as $\mathcal{CW}(\boldsymbol{\Sigma}, M)$, it can be proved that the Schur complement $\boldsymbol{W}/\boldsymbol{C}$ is also complex-Wishart distributed, as $\mathcal{CW}(\boldsymbol{\Sigma}/\boldsymbol{E}, M-r)$, where $\boldsymbol{E}$ is the lower-right $r \times r$ block in $\boldsymbol{\Sigma}$, analogous to $\boldsymbol{C}$. References [58, 59] provide proof of this in the real case, i.e. when $\boldsymbol{W}$ is real-Wishart distributed. However, we checked numerically that the same result holds in the complex case too. We checked this with

various values for M and r.
We highlight that, in case $\boldsymbol{A}$ is one-dimensional, then its complex Wishart distribution reduces to a 1-D Gamma distribution. Additionally, in that case, the Schur complement $\boldsymbol{\Sigma}/\boldsymbol{E}$ corresponds to $1/(\boldsymbol{\Sigma}^{-1})_{11}$. Finally,

$$\mathcal{CW}\left(\boldsymbol{\Sigma}/\boldsymbol{S}, M-r\right)_{\text{1-D}} = \Gamma\left(M-r, M\left(\boldsymbol{\Sigma}^{-1}\right)_{11}\right) \tag{3.16}$$

We note that this decorrelation holds only in case $M > r$, because of the definition of Gamma distribution.

3.3 Bayesian inference of spectral parameters

In the previous section, we discussed the distributions ruling the experimental outcome of Gaussian stochastic processes. Given a random noisy process governed by a PSD S, or more generically by a CPSD matrix $\boldsymbol{\Sigma}$, those distributions allow estimating the probability of certain experimental outcomes. In the real experimental case, though, one is usually interested in the inverse process, that is, understanding the underlying theoretical parameters on the basis of the observed experimental outcome. Hence, what one would really be interested in looking for is the probability distribution of the theoretical parameters, given a set of experimental values. The inversion of point of view requires changing the approach, and shifting to a Bayesian analysis. This change of perspective allows us to define distributions of the theoretical spectral parameters, based on observations, and finally to infer their uncertainties.

In the Bayesian language, this means that our purpose is to estimate the theoretical parameters θ, conditional to the measured experimental samples x. It follows from Bayes' theorem that

$$p(\boldsymbol{\theta}|\boldsymbol{x}) = \frac{p(\boldsymbol{x}|\boldsymbol{\theta})\, p(\boldsymbol{\theta})}{p(\boldsymbol{x})}, \tag{3.17}$$

here, $p(\boldsymbol{\theta}|\boldsymbol{x})$ is known as the posterior probability for $\boldsymbol{\theta}$, and $p(\boldsymbol{x}|\boldsymbol{\theta})$ is the distribution of the experimental measurements, conditional to the underlying unknown parameters $\boldsymbol{\theta}$. The term $p(\boldsymbol{x})$ is just a normalization factor.

What requires a deeper analysis is the term $p(\boldsymbol{\theta})$, which is the so-called prior distribution, any *a priori* assumptions on the unknown parameters $p(\boldsymbol{\theta})$, that we deal with in the following paragraphs.

3.3.1 Choice of Bayesian prior

The choice of the prior is a delicate and well-known issue in Bayesian analysis. It needs to be carefully addressed, as it can heavily affect the result of the posterior probability distribution.

Some prior choices are rather obvious:

1. In case the parameters $\boldsymbol{\theta}$ are subject to some constraints for physical reasons, which are not taken into account by the experimental observation $p(\boldsymbol{x}|\boldsymbol{\theta})$. Including such constraints is essential to ensure a realistic result.

2. In case there is some additional experimental evidence on the probability distribution of the parameters. The unknown parameters could indeed be physical properties of the experimental system, and as such they could have been analyzed by different experiments. This is roughly the case in the LPF analysis: some parameters, as for example the stiffnesses correlating the TM positioning to their acceleration, do have prior knowledge from system identification experiments. In that case, the prior distribution would be so narrow that it would behave as a constraint for the parameters $\boldsymbol{\theta}$.

3. The most probable case is that in which one does not have any prior assumption on the unknown parameters, and they want to be as agnostic as possible. This is the case in which one needs the class of the non-informative priors. It has been thoroughly addressed and it's currently an active field of research. One could naively assume that choosing a flat prior (i.e., neglecting $p(\boldsymbol{\theta})$ in Eq. (3.17)) could be a good non-informative choice. It indeed is, in a large family of cases, but it's not always the best choice, since a simple reparametrization of the system would lead to a completely different prior assumption. Jeffrey addressed this issue by proposing the well-known Jeffrey's non-informative prior [60], which is invariant under a change of coordinates.

   In case a parameter θ is a scale parameter – which is indeed the case of the PSD distribution –, Jeffrey's non-informative prior is $1/\theta$, which corresponds to choosing a flat prior on $\log\theta$.

Matrix formulation: posterior distribution

A possible straightforward approach would be to find directly a posterior for the whole $\boldsymbol{\Sigma}$ matrix, which would obviously depend only on the measured $\boldsymbol{W}$, as $p(\boldsymbol{\Sigma}|\boldsymbol{W})$. This approach would yield some advantages: it would allow to have the whole joint posterior distribution of all the matrix elements; this would consequently allow to deduce the marginal posterior distribution for all the spectral quantities of interest, either analytically or numerically.

However, the calculation of such a posterior would require the choice of a suitable prior, which would be a prior on the whole matrix $\boldsymbol{\Sigma}$. This makes the computation a little more complicated, and the results have some limitations that we now analyze:

1. Defining the inverse-matrix $\boldsymbol{Q} = \boldsymbol{\Sigma}^{-1}$, it can be noted that it is on its turn distributed as a complex Wishart random variable, with $M+p$ degrees of freedom: $\boldsymbol{Q} \sim \mathcal{CW}(\boldsymbol{W}^{-1}, M+p)$. In case of a flat prior on the inverse CPSD matrix $\boldsymbol{Q}$, that is $p(\boldsymbol{\Sigma}) \propto 1$, the distribution $p(\boldsymbol{W}|\boldsymbol{\Sigma})$ would then be a complete posterior for $\boldsymbol{\Sigma}$, that is $p(\boldsymbol{\Sigma}|\boldsymbol{W}) = p(\boldsymbol{W}|\boldsymbol{\Sigma})$. Therefore, samples of the inverse-posterior matrix $\boldsymbol{Q}$ could be drawn using the same random generator as described in Appendix B.1, and samples of $\boldsymbol{\Sigma}$ could be found by just inverting the matrix. We note, following [61], that this case would yield no bias on the posterior for $\boldsymbol{\Sigma}$, as the expected value would be $\langle\boldsymbol{\Sigma}\rangle = \boldsymbol{W}$.

2. As discussed, a sounder approach would be that of Jeffrey's non-informative prior. In our case, Jeffrey's prior for $\boldsymbol{Q}$ can be computed and is proportional to

$|\boldsymbol{Q}|^{-p}$ [62]. Implementing this choice, the probability distribution of $\boldsymbol{Q}$ would result $\boldsymbol{Q} \sim \mathcal{CW}(\boldsymbol{W}^{-1}, M)$. Even though this approach is better for being non-informative, the posterior of $\boldsymbol{\Sigma}$ is not unbiased: $\langle \boldsymbol{\Sigma} \rangle = \frac{M}{M-p}\boldsymbol{W}$.

It is clear that this represents a limitation of this prior choice: the resulting marginal posteriors of the PSDs show a dependence on the dimension p of the CPSD matrix. This can be seen as a dependence of the individual PSDs on the number of measurements performed, which is unphysical. We have checked this behavior numerically, checking the distributions of the diagonal elements of randomly-extracted $\boldsymbol{\Sigma}$ matrices.

We note, though, that the discrepancy is significant only if the number of periodograms is very low.

In the following, we propose posteriors for the theoretical PSDs and MSCs, which are only based on the experimental value of the corresponding variable.

3.3.2 PSD posterior inference

A relevant case is undoubtedly the one-dimensional one. This case corresponds to the measurement of a single time series. The CPSD matrix $\boldsymbol{\Sigma}$ reduces to S, the theoretical PSD of the time series. Let's assume that M independent periodograms have been measured, and that an experimental PSD Π has been found.

We apply Bayes' theorem, to find the conditional probability of the theoretical S, based on the measurement Π. We choose Jeffrey's prior. As already said, as S is a scale parameter, this means a uniform prior in $\log S$, i.e., $p(S) \propto 1/S$.

Given these facts, finally, the posterior distribution of S follows an inverse-gamma distribution with M degrees of freedom:

$$p\left(S|\Pi\right) \sim \text{invΓ}\left(M, M\Pi\right) \tag{3.18}$$

We note that the mean value of the $\text{invΓ}\left(M, M\Pi\right)$ distribution has a slight bias relative to the observed value: $\langle S \rangle = \Pi\frac{M}{M-1}$, which vanishes when using a large number of averaging windows. It is interesting to note that this distribution is well-defined also if $M = 1$, thanks to Jeffrey's prior: however, in that case it is not possible to define the mean value.

Based on this posterior distribution, we compute the equally-tailed p_v-credible interval from the $(1 \pm p_v)/2$ quantiles of the distribution. Finally, we can give an estimate of the errors on S, at a given confidence level p_v:

$$S_{\pm} = \frac{M\Pi}{\Gamma_u^{-1}\left(M, (1 \pm p_v)/2\right)} \tag{3.19}$$

here, Γ_u^{-1} is the inverse upper incomplete Gamma function. Usually, p_v corresponds to 68.3%, 95.5%, 99.7%, respectively corresponding to the 1σ, 2σ, 3σ confidence intervals of the normal distribution. In Fig. 3.3, we provide an example of posterior PSD distribution, with the associated mean value and confidence intervals.

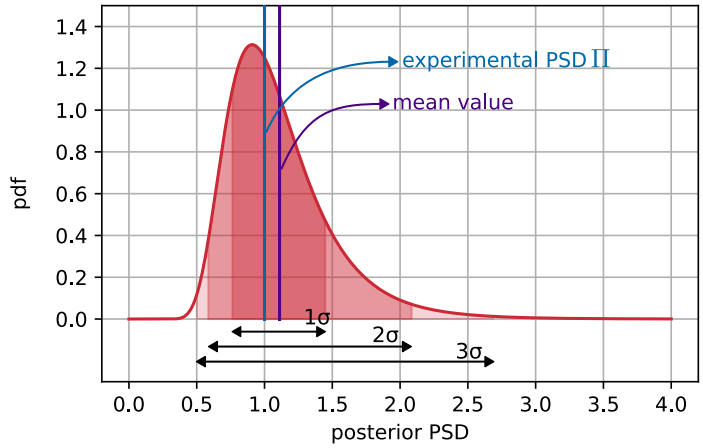


Figure 3.3: Example of posterior PSD distributions of the theoretical PSD S, given an experimental PSD $\Pi = 1$, and a number of periodograms $M = 10$. We highlight the 1σ, 2σ, 3σ confidence intervals, and show that the distribution is slightly biased, as described in text.

3.3.3 Cross-coherence and multiple coherence, posterior inference

Along the same line, we use Eq. (3.12) to derive a posterior distribution for the theoretical MSC $|\rho_{ij}|^2$, and the multiple coherence R^2. We employ flat prior distributions, not to add any additional constraints. For the theoretical MSC, Eq. (3.12) yields:

$$p(|\rho|^2||\hat{\rho}|^2) = (M+1)(1-|\rho|^2)^M \times \frac{{}_2F_1(M,M,1,|\hat{\rho}|^2\,|\rho|^2)}{{}_2F_1(M,M,2+M,|\hat{\rho}|^2)} \tag{3.20}$$

Similarly, for R^2, Eq. (3.13) yields:

$$p(R^2|\hat{R}^2) = (M+1)\,{}_2F_1(M,M,p-1,\hat{R}^2R^2) \times(1-R^2)^M/{}_pF_q\left(\begin{matrix}(1,M,M)\\(M+2,p-1)\end{matrix};\hat{R}^2\right) \tag{3.21}$$

where ${}_pF_q$ is the generalized hypergeometric function.

In Table 3.1, we provide a summary of the probability distribution analyzed in this section: the sample (experimental) PSD and MSC, conditional to the theoretical values; the theoretical PSD, MSC, and multiple coherence, conditional to the experimental measurement.

Table 3.1: Summary of the probability distributions of the sample PSD and MSC, and theoretical PSD, MSC and multiple coherence. For the sake of clarity, the latter are not normalized.

Sample PSD, Eq. (3.11)	$p\left(\Pi\middle\vert S\right) \sim \Gamma\left(\alpha = M, \beta = M/S\right)$
Sample MSC, Eq. (3.12)	$p(\lvert\hat{\rho}\rvert^2\vert\rho\rvert^2) \propto (1-\lvert\hat{\rho}\rvert^2)^{M-2}\,{}_2F_1(M, M, 1, \lvert\hat{\rho}\rvert^2\lvert\rho\rvert^2)$
Th. PSD, Eq. (3.18)	$p\left(S\middle\vert\Pi\right) \sim \mathrm{inv}\Gamma\left(\alpha = M, \beta = M\Pi\right)$
Th. MSC, Eq. (3.20)	$p(\lvert\rho\rvert^2\vert\lvert\hat{\rho}\rvert^2) \propto (1-\lvert\rho\rvert^2)^{M}\,{}_2F_1(M, M, 1, \lvert\hat{\rho}\rvert^2\lvert\rho\rvert^2)$
Th. Mult. coherence, Eq. (3.21)	$p(R^2\vert\hat{R}^2) \propto (1-R^2)^M {}_2F_1(M, M, p-1, \hat{R}^2R^2)$

3.4 Time series decorrelation: modelization and single-frequency case

The findings of Sec. 3.3 are also at the basis of our decorrelation procedure, whose ultimate purpose is to analyze and subtract the contribution of a time series to the noise of another one. In the LPF case, this tool is extremely useful for the estimation of the contribution to the acceleration noise (Chapter 6). The procedure needs all the time series (the main one and the ones to be subtracted) to be simultaneously measured, and to be synchronously sampled. We employ the following scheme:

1. We first analyze the single-frequency case, at a given window length, and a fixed number of independent periodograms. We find that this case is strictly related to the quantities analyzed in Sec. 3.3, and that it is possible to give the results in analytical form.

2. We analyze the correlation between adjacent frequencies, and employ a scheme that allows us to minimize it.

3. After finding quasi-independent frequencies, we extend the method to the multi-frequency case, allowing signal decorrelation in chosen frequency bands. We employ numerical methods to analyze the modeled probability distribution.

4. In Chapter 6, we describe how we applied this procedure for LPF analyses [5]. This procedure has also been implemented for the analyses of the LPF actuation noise [23].

This procedure extends the findings of [51], relying on the same modelization.

3.4.1 Model definition

Let's use the LPF nomenclature, though this method is fully general. Let's assume that the main measurement $\Delta g(t)$ is contributed by a number of time series $z_i(t)$, $i = 1, \ldots, r$, and by an independent term $\Delta g_0(t)$, which we call the *residual* noise. The

contributing signals $z_i(t)$ couple linearly to $\Delta g(t)$ through the coupling coefficients α_i. We refer to these coefficients as the *susceptibilities*, and assume that they are real-valued and time-independent. In the general case, the signals $z_i(t)$ could also be cross-correlating with one another, could be phase-shifted, and could be affected by some additional independent readout noise.

All in all, the model reads:

$$\begin{cases} \Delta g(t) &= \Delta g_0(t) + \sum_{i=1}^{r} \alpha_i z_i(t) \\ y_i(t) &= z_i(t) + n_i(t) \end{cases} \tag{3.22}$$

where we have introduced $n_i(t)$, the readout noise affecting the i-th signal time series $y_i(t)$, and $z_i(t)$, the "true" signal. The coupling coefficients α_i are generally considered time-independent, at least on a single LPF noise run. Experimentally, only Δg and y_i are measured. The purpose of the decorrelation procedure is to analyze to which extent the y_i contribute to Δg, and to infer the (complex) susceptibilities α_i.
The model in Eq. (3.22) is easily readable, however not complete. In the general case, the susceptibilities α_i could depend on time, including a delay between the time series $z_i(t)$ and its effect on $\Delta g(t)$, behaving as a linear filter. A clear example is that of thermal effects, which are characterized by a slow time scale: as temperature varies, the effect of its variation is detected a lot of time later, delayed with respect to the variation itself. The general model reads

$$\begin{cases} \Delta g(t) &= \Delta g_0(t) + \sum_{i=1}^{r} \int_{-\infty}^{+\infty} \alpha_i(t-t') z_i(t)\, dt' \\ y_i(t) &= z_i(t) + n_i(t) \end{cases} \tag{3.23}$$

As we will perform analyses in the frequency domain, we will generally deal with the (complex) Fourier transform of the αs. Thanks to the properties of the Fourier transform, the convolution in the first line in Eq. (3.23) can be rewritten as a simple product in the frequency domain. One can see that, treating $\Delta g(t)$ and $n_i(t), z_i(t)$ as stationary-noise time series, the PSDs of terms in Eq. (3.23) read

$$\begin{cases} S_{\Delta g \Delta g}(\omega) &= S_{\Delta g_0}(\omega) + \sum_{i,j} \alpha_i(\omega) S_{z_i z_j}(\omega) \alpha_j^*(\omega) \\ S_{y_i y_i}(\omega) &= S_{n_i n_i}(\omega) + S_{z_i z_i}(\omega) \end{cases} \tag{3.24}$$

Calculations leading to the results in Eq. (3.24) are postponed to App. B.2.1. In case α is time-independent, the dependence of α_i on the frequency ω drops in Eq. (3.22), so that the model reads

$$\begin{cases} S_{\Delta g \Delta g}(\omega) &= S_{\Delta g_0}(\omega) + \sum_{i,j} \alpha_i S_{z_i z_j}(\omega) \alpha_j \\ S_{y_i y_i}(\omega) &= S_{n_i n_i}(\omega) + S_{z_i z_i}(\omega) \end{cases} \tag{3.25}$$

However, this result can be more easily manipulated in matrix form, casting Δg and y_i as a CPSD matrix, which we will manipulate with the results of the previous sections.

Let's take the frequency f as reference, and assume that M overlapping periodograms are available. We consider as array of measurements $\{\Delta g, \boldsymbol{y}\}$. The sample

CPSD, $\mathbf{\Pi}$, is therefore the $(r+1)\times(r+1)$ matrix following the complex Wishart distribution $\boldsymbol{W} = M\mathbf{\Pi} \sim \mathcal{CW}(\mathbf{\Sigma}, M)$. For clarity, we decompose $\mathbf{\Sigma}$ and $\mathbf{\Pi}$ into blocks, as in Eq. (3.26). We denote with $P, \mathbf{\Pi}$ the sample values, and with $S, \mathbf{\Sigma}$ the theoretical values.

$$\mathbf{\Sigma} = \left(\begin{array}{c|c} S_{\Delta g \Delta g} & \boldsymbol{S}_{\Delta g y} \\ \hline \boldsymbol{S}^H_{\Delta g y} & \boldsymbol{S}_{yy} \end{array}\right) \qquad \mathbf{\Pi} = \left(\begin{array}{c|c} P_{\Delta g \Delta g} & \boldsymbol{P}_{\Delta g y} \\ \hline \boldsymbol{P}^H_{\Delta g y} & \boldsymbol{P}_{yy} \end{array}\right) \tag{3.26}$$

Observation on readout noise. Before proceeding, we make an observation about the inclusion of readout noise n_i in the model in Eq. (3.25). We note that including the readout noise would lead to an overparametrization of the model.
Let's focus on the matrix form as in Eq. (3.26), in the simple case with only one signal y. the free parameters in the model are: $S_{\Delta g_0}(f)$, $S_z(f)$, $S_n(f)$, α. Summing up, assuming a total number of frequencies N_f, the total number of parameters is $1+3N_f$. At the same time, the number of free elements (data) in the Hermitian matrix is $3N_f$. This means that the model is overparameterized, as the number of parameters exceeds the available data. In case the matrix dimension is higher, the issue is not overcome.

Therefore, we proceed as follows:

1. First, we analyze the case in which the readout noise is negligible in the analyzed frequency band. In that case, we assume that the susceptibilities are not known, and have to be recovered with the decorrelation method.

2. Later, we analyze which is the effect of non-negligible readout noise on contribution estimation. In other words, we analyze to what extent the resulting distributions are biased by the presence of unmodeled noise affecting the signals $z_i(t)$, and how that problem can be addressed.

3.4.2 Single frequency time series decorrelation

We proceed considering that the readout noise is negligible: $y_i(t) \sim z_i(t)$. Given the model in Eq. (3.22), (3.23) and the results thereafter, we also show the connection between the spectral quantities defined in Sec. 3.2 and quantities of interest for signal decorrelation. Mathematical details and derivations of the following results are in App. B.2.2.

1. The upper-right line vector, representing the cross-correlations between Δg and the signals y, can be expressed as

$$\boldsymbol{S}_{\Delta g y} = \boldsymbol{\alpha} \cdot \boldsymbol{S}_{yy}, \tag{3.27}$$

   where $\boldsymbol{\alpha}$ is the vector of the complex susceptibilities. Their magnitudes are the magnitudes of the real susceptibilities in Eq. (3.22).

2. In case time series are correlated with a time delay τ_i, such information is gathered by the phase-shift of the susceptibilities, ϕ_i, through $\phi_i = 2\pi f \tau_i$.

Assume that the effect on $\Delta g(t)$ of the time series $y(t)$ is delayed by a time τ, i.e., $\Delta g(t) = \alpha\, y(t-\tau)$, with α a real number. According to the CPSD definition and estimation in Eq. (3.5), and the definition of DFT transform in Eq. 3.3, it is clear that the CPSD between $\Delta g(t)$ and $y(t)$ is

$$S_{\Delta g\, y}(f) = \alpha e^{2i\pi f\tau} S_{yy}, \tag{3.28}$$

meaning that the complex susceptibility carries a phase shift $\phi = 2\pi f\tau$, related to the time delay.

3. The susceptibilities can easily be linked to the inverse matrix $\boldsymbol{Q} = \boldsymbol{\Sigma}^{-1}$. It can indeed be proved (see App. B.2.2) that

$$\alpha_j = -\frac{Q_{1,j+1}}{Q_{11}} \tag{3.29}$$

4. The total contribution to noise by the y_i, and the residual noise, can also be expanded:

$$S_{\text{contr}} = \boldsymbol{\alpha} \cdot \boldsymbol{S}_{yy} \cdot \boldsymbol{\alpha}^H = \boldsymbol{S}_{\Delta gy} \cdot \boldsymbol{S}_{yy}^{-1} \cdot \boldsymbol{S}_{\Delta gy}^H \tag{3.30}$$

$$S_{\text{res}} = S_{\Delta g\Delta g} - S_{\text{contr}} = \frac{1}{Q_{11}} \tag{3.31}$$

5. The residual noise coincides with the Schur complement of the matrix $\boldsymbol{S}_{yy}$ in $\boldsymbol{\Sigma}$.

6. The power contribution fraction of the signals y_i to Δg, $S_{\text{contr}}/S_{\Delta g\Delta g}$, coincides with the multiple coherence R^2.

3.4.3 Estimation of the residual noise: single-frequency case

It is now clear how the CPSD matrix formulation can be used for signal decorrelation, as the sample residual noise coincides with the Schur complement of the sample signal matrix $\boldsymbol{P}_{yy}$ in the overall CPSD matrix $\boldsymbol{\Pi}$,

$$P_{\text{res}} = \boldsymbol{\Pi}/\boldsymbol{P}_{yy} = P_{\Delta g\Delta g} - \boldsymbol{P}_{\Delta gy} \cdot \boldsymbol{P}_{yy}^{-1} \cdot \boldsymbol{P}_{\Delta gy}^H \tag{3.32}$$

In Sec. 3.2.5, we also noted that the sample residual noise is distributed as a complex Wishart distribution. Indeed, since $M\boldsymbol{\Pi}$ is complex Wishart distributed, and $M\boldsymbol{P}_{yy}$ is a $r \times r$ matrix, then the sample residual noise is distributed as a complex Wishart, with $M - r$ degrees of freedom:

$$P_{\text{res}} \sim \mathcal{CW}\left(\boldsymbol{\Sigma}/\boldsymbol{S}_{yy}, M - r\right), \tag{3.33}$$

where $\boldsymbol{\Sigma}/\boldsymbol{S}_{yy}$ is the Schur complement of $\boldsymbol{S}_{yy}$ in $\boldsymbol{\Sigma}$, *i.e.*, the underlying theoretical residual noise.

We note that this result reduces, in one dimension (Eq. (3.11)), to the Gamma distribution $P_{\text{res}} \sim \Gamma\left(M - r, (\boldsymbol{\Sigma}/M\boldsymbol{S}_{yy})^{-1}\right)$.

The corresponding posterior distribution of the theoretical residual noise, under the assumption of a noninformative prior $1/S_{\rm res}$, is therefore

$$S_{\rm res} \sim \text{invΓ}\left(M - r, M\,\mathbf{\Pi}/\boldsymbol{P}_{yy}\right) \tag{3.34}$$

We note that the noise decorrelation can be performed only in case the number of periodograms is greater than the total number of decorrelated signals, *i.e.*, $M > r$.

We do not find an explicit analytical expression for the probability distribution of the real/imaginary parts of the susceptibility. However, we refer the reader to Sec. 3.6.3 for numerical tests: the single-frequency case can be tested numerically with the tools proposed for the multi-frequency case. In Fig. 3.6 on page 69, we show a comparison along this line, between the single-frequency analytical model in Eq. (3.34) and the multi-frequency numerical model with one frequency only.

3.5 Choice of spectral frequencies

It is widely known that the commonly used Welch method for the estimation of (cross-)PSDs has limitations in its application. What is usually noted is that the choice of the periodogram length is arbitrary and non-optimized: estimates at high frequencies are computed with fewer periodograms than available, leading to non-optimized variance [51, 63]. Moreover, the same references note that spectra are usually analyzed in a logarithmically-spaced fashion, whilst the Welch method provides linearly-spaced points. Increasing the number of averaged periodograms at high frequencies is paramount to reducing the variance in the estimates, as it is indeed done in [51, 63].Here, we face another issue: in the next section, we face the need to combine single-frequency probability distributions, in order to obtain multi-frequency joint distributions. Hence, we need them to be as independent as possible. We follow the frequency choice also used in [2, 3]: the scheme allows to find quasi-independent frequencies, logarithmically spaced. The spacing between the frequencies is such that f_{i+1}/f_i is roughly 5/3, from the second bin onwards, choosing for each periodogram the 9th spectral index.

i	f_i(mHz)	i	f_i(mHz)	i	f_i(mHz)
1	0.018	8	0.77	15	28.
2	0.036	9	1.3	16	46.
3	0.060	10	2.1	17	77.
4	0.10	11	3.6	18	128.
5	0.17	12	6.0	19	213.
6	0.28	13	9.9	20	354.
7	0.46	14	17.	21	591.
				22	985.

Table 3.2: List of frequencies used for LPF spectral estimation, in the band 18 µHz–1 Hz. The lowest frequency is chosen so that the 4th frequency is exactly 0.1 mHz, the lower boundary of the LISA band.

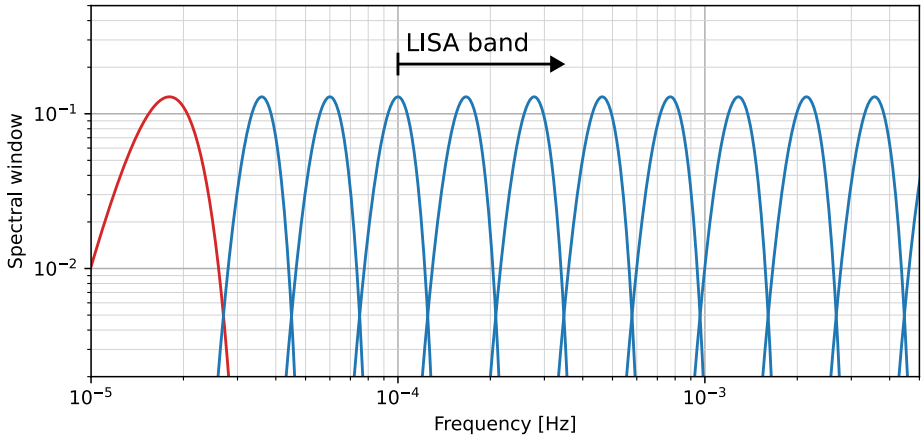


Figure 3.4: Overlap of spectral windows (4-term Blackman-Harris −92 dB, scaled to the same height) evaluated at logarithmically-spaced frequencies, following the optimized scheme in text. We note that the first frequency bin (18 µHz in the LPF case) is wider than the following ones, ad thus contains a larger pickup from lower frequencies.

With this choice, the correlation among adjacent (C)PSDs is below 5%. The first frequency bin is chosen differently, using the same periodogram set as for the second bin, picking the 5th spectral index instead of the 9th. This tradeoff allows to analyze frequencies down to 18 µHz, while keeping the correlation to the DC term at about 3%. In Fig. 3.4, we show the overlap among the frequency windows used for LPF analysis, up to 5 mHz. As discussed, we note that the first bin is wider because of its definition.

In Table 3.2, we list the frequency employed for the LPF analyses, in the band 18 µHz–1 Hz. The lowest frequency is chosen so that one of the frequencies is exactly 0.1 mHz. Referring to the noise run Feb17B, the number of periodograms used in calculations spans from 9 at 18 µHz to 287903 at 0.985 Hz.

3.6 Multi-frequency time series decorrelation

In the LPF case, the susceptibilities in the model Eq. (3.22) are constant, fixed in time. Hence, their Fourier transforms do not depend on frequency. The modelization proposed in Sec. 3.4 can then be extended to a set of frequencies, developing a model which is valid on a wide frequency band.

Here we propose a multi-frequency method, that allows to decorrelate a set of signals from a main time series, on a selected set of frequencies. As in the single-frequency case, we assume that the readout noise $n_i(t)$ is negligible on the frequency band under analysis, and hence $y_i(t) \sim z_i(t)$.

Let's start with the distribution of the CPSD matrices at each frequency. If frequencies are chosen to ensure minimal correlation, as described in Sec. 3.5, it is straightforward to see that the joint distribution of the (sample) CPSD set is simply

given by multiplication of the individual complex Wishart distributions[3],

$$p(\{\boldsymbol{W}\}|\{\boldsymbol{\Sigma}\}) = \prod_f \left(\frac{|\boldsymbol{W}_f|^{M_f - p}}{\widetilde{\Gamma}_p(M_f)|\boldsymbol{\Sigma}_f|^{M_f}} \operatorname{etr}\left[-\boldsymbol{\Sigma}_f^{-1}\boldsymbol{W}_f\right] \right). \tag{3.35}$$

Here, the subscript f explicates the dependence on frequency. Moreover, we note that one can re-write $\boldsymbol{\Sigma}$ as

$$\boldsymbol{\Sigma}(f) = \left(\begin{array}{c|c} S_{\Delta g_0}(f) + \boldsymbol{\alpha} \cdot \boldsymbol{S}_{yy}(f) \cdot \boldsymbol{\alpha}^H & \boldsymbol{\alpha} \cdot \boldsymbol{S}_{yy}(f) \\ \hline \boldsymbol{S}_{yy}^H(f) \cdot \boldsymbol{\alpha}^H & \boldsymbol{S}_{yy}(f) \end{array} \right) \tag{3.36}$$

With this modelization, one may proceed applying Bayes' theorem to find the posterior distribution of the theoretical parameters. However, the application is not as straightforward as it is in the single-frequency case, and more discussion is needed.

Tentative direct – inaccurate – approach

One could be tempted to approach this problem naively, expressing $\boldsymbol{\Sigma}$ with the unknown parameters and, later, employing numerical methods to find the marginal posteriors. For instance, one could re-write $\boldsymbol{\Sigma}$ as in Eq. (3.36), and look for a Bayesian posterior as

$$p(\boldsymbol{\alpha}, \boldsymbol{S}_{yy}(f), S_{\Delta g_0}(f) \,|\, \boldsymbol{W}) = \frac{p(\boldsymbol{W} \,|\, \boldsymbol{\Sigma})}{p(\boldsymbol{W})}\, p(\boldsymbol{\alpha}, \boldsymbol{S}_{yy}(f), S_{\Delta g_0}(f)) \tag{3.37}$$

We identify a major issue in this approach that prevents its use as is, requiring further analysis.

The issue that we identify regards the properties of the $\boldsymbol{\Sigma}$ matrix itself. As stated in Sec. 3.2, the theoretical CPSD matrix $\boldsymbol{\Sigma}$ is positive-definite. However, as it is defined in Eq. (3.36), this property is not mathematically ensured. Approaching the problem with numerical methods, the theoretical parameters $\boldsymbol{\alpha}, \boldsymbol{S}_{yy}(f), S_{\Delta g_0}(f)$ could assume values that make $\boldsymbol{\Sigma}$ not positive definite, yielding a negative determinant, and making the computation crash.

One could include this constraint as a hard cut in the prior distribution. However, the probability of randomly generating a positive-definite matrix decreases as the matrix dimension increases, making it even harder and harder for random samples to be in the correct domain. Sampling high-dimensional matrices would then require a lot of computational time, most of which wasted in looking for positive-definite samples.

Changing coordinates to ensure that $\boldsymbol{\Sigma}$ is positive-definite would be even harder, requiring non-linear changes of variables, hard to handle.

[3]Provided that $M_f \geq p$ for all frequencies f, otherwise some of the terms would be degenerate. We will address this issue later.

3.6.1 Estimate of susceptibilities and residuals

In order to overcome the identified issue, we elaborate on the properties of the CPSD matrices, and the complex Wishart distributions. We work on the single-frequency distribution, aiming at expanding to the multi-frequency case. First, let's assume that the number of available periodograms M at frequency f is greater than or equal to the number of time series p, so that the CPSD matrix follows the (non-singular) complex Wishart matrix.
We re-write the complex Wishart distribution, for simplicity:

$$p(\boldsymbol{W}|\boldsymbol{\Sigma}, M) = \frac{|\boldsymbol{W}|^{M-p}}{\widetilde{\Gamma}_p(M)|\boldsymbol{\Sigma}|^M} \operatorname{etr}\left[-\boldsymbol{\Sigma}^{-1}\boldsymbol{W}\right]$$

In addition to this, we assume the modelization of the $\boldsymbol{\Sigma}$ matrix in Eq. (3.36). We follow some steps that lead us to a manipulation of the probability distribution function, so that it is more easily used.

1. First, we define the matrix $\boldsymbol{U}$ and the vector $\boldsymbol{\eta}$, as follows:

$$\boldsymbol{U} = \begin{pmatrix} 1 & \boldsymbol{\alpha} \\ \boldsymbol{0} & \mathbb{I} \end{pmatrix} \qquad \boldsymbol{\eta} = (-1, \boldsymbol{\alpha}) \tag{3.38}$$

   Where $\mathbb{I}$ is the $(p-1)\times(p-1)$ identity matrix.
   The matrix $\boldsymbol{U}$ represents a linear transformation $\Delta g_0 \to \Delta g$, $y_i \to y_y$. We note that the inverse $\boldsymbol{U}^{-1}$ is very similar to $\boldsymbol{U}$ itself, with $\boldsymbol{\alpha} \to -\boldsymbol{\alpha}$. The definition of $\boldsymbol{U}$ leads to a first simplification: as $\Delta g_0(t)$ is independent of $\boldsymbol{y}(t)$, the processes after the linear transformation can be represented with a block matrix,

$$\boldsymbol{\Sigma}' = \begin{pmatrix} S_{\Delta g_0} & \\ & \boldsymbol{S}_{yy} \end{pmatrix}, \qquad \boldsymbol{\Sigma} = \boldsymbol{U}\boldsymbol{\Sigma}'\boldsymbol{U}^H \tag{3.39}$$

   Notice that $\boldsymbol{U}$ is real, so it makes no sense to write $(\cdot)^H$ instead of $(\cdot)^T$. However, we still use $(\cdot)^H$ since some other quantities are complex.

2. We aim at replacing $\boldsymbol{\Sigma}$ with $\boldsymbol{\Sigma}'$ in the probability distribution. Hence, we first note that the determinant of $\boldsymbol{U}$ is 1, and that the determinant of $\boldsymbol{\Sigma}$ can be simply re-written as

$$|\boldsymbol{\Sigma}'| = |\boldsymbol{U}|\,|\boldsymbol{\Sigma}|'\,|\boldsymbol{U}^H| = |\boldsymbol{\Sigma}|' = S_{\Delta g_0}\,|\boldsymbol{S}_{yy}| \tag{3.40}$$

3. The exponent can be rewritten as well: let's note that

$$\begin{aligned} \operatorname{etr}\left[-\boldsymbol{\Sigma}^{-1}\boldsymbol{W}\right] &= \operatorname{etr}\left[-\boldsymbol{U}^{-H}\boldsymbol{\Sigma}'^{-1}\boldsymbol{U}^{-1}\boldsymbol{W}\right] = \\ &= \operatorname{etr}\left[-\boldsymbol{\Sigma}'^{-1}\boldsymbol{U}^{-1}\boldsymbol{W}\boldsymbol{U}^{-H}\right] = \operatorname{etr}\left[-\boldsymbol{\Sigma}'^{-1}\boldsymbol{W}'\right], \end{aligned} \tag{3.41}$$

where $\boldsymbol{W}' = \boldsymbol{U}^{-1}\boldsymbol{W}\boldsymbol{U}^{-H}$, which can be further elaborated,

$$\begin{aligned}\boldsymbol{W}' = \boldsymbol{U}^{-1}\boldsymbol{W}\boldsymbol{U}^{-H} &= \begin{pmatrix} \boldsymbol{W}'_{11} & \boldsymbol{W}'_{1y} \\ W'^{H}_{1y} & \boldsymbol{W}'_{yy} \end{pmatrix} = \\ &= \begin{pmatrix} W_{11} - \boldsymbol{W}_{1y}\boldsymbol{\alpha}^H - \boldsymbol{\alpha}\boldsymbol{W}^H_{1y} + \boldsymbol{\alpha}\boldsymbol{W}_{yy}\boldsymbol{\alpha}^H & \boldsymbol{W}_{1y} - \boldsymbol{\alpha}\boldsymbol{W}_{yy} \\ \boldsymbol{W}^H_{1y} - \boldsymbol{W}_{yy}\boldsymbol{\alpha}^H & \boldsymbol{W}_{yy} \end{pmatrix} = \\ &= \begin{pmatrix} \boldsymbol{\eta}\boldsymbol{W}\boldsymbol{\eta}^H & \dots \\ \dots & \boldsymbol{W}_{yy} \end{pmatrix}\end{aligned} \tag{3.42}$$

The dots indicate something we are not interested in. In the last step, we note that $\boldsymbol{W}'_{11}$ can be easily rewritten in terms of $\boldsymbol{\eta}$ and $\boldsymbol{W}$, and $\boldsymbol{W}'_{yy}$ is just $\boldsymbol{W}_{yy}$.

One can also easily calculate the inverse of $\boldsymbol{\Sigma}'$, which turns out to be useful as it multiplies $\boldsymbol{W}'$ in the likelihood function.

$$\boldsymbol{\Sigma}'^{-1} = \begin{pmatrix} 1/S_{\Delta g_0} & \\ & \boldsymbol{S}^{-1}_{yy} \end{pmatrix} \tag{3.43}$$

4. All in all, the conclusion on the probability density function (still single-frequency) is

$$\begin{aligned} p(\boldsymbol{W}|\boldsymbol{\alpha}, S_{\Delta g_0}, \boldsymbol{\Sigma}_{yy}) &\propto \frac{1}{|\boldsymbol{\Sigma}'|^M} \operatorname{etr}\left[-\boldsymbol{\Sigma}'^{-1}\boldsymbol{W}'\right] \propto \\ &\propto \frac{1}{S^M_{\Delta g_0}} \exp\left[-\frac{\boldsymbol{\eta}\boldsymbol{W}\boldsymbol{\eta}^H}{S_{\Delta g_0}}\right] \frac{1}{|\boldsymbol{S}_{yy}|^M} \operatorname{etr}\left[-\boldsymbol{S}^{-1}_{yy}\boldsymbol{W}_{yy}\right] \end{aligned} \tag{3.44}$$

This expression can be easily exported to the multi-frequency case but, before doing that, we notice an additional point. The matrix $\boldsymbol{S}_{yy}$ is the posterior matrix of the signals $\boldsymbol{y}(t)$, for which one may impose assumptions. One reasonable hypothesis is that $\boldsymbol{S}_{yy}$ can be decomposed as a block-diagonal matrix, grouping subsets of processes that cross-correlate. For instance, assuming that it can be decomposed into k blocks,

$$\boldsymbol{S}_{yy} = \begin{pmatrix} \boldsymbol{S}_{yy,1} & & \\ & \ddots & \\ & & \boldsymbol{S}_{yy,k} \end{pmatrix} \tag{3.45}$$

This allows us to rewrite the exponential term, yielding a final expression that separates the independent blocks.

$$\frac{1}{|\boldsymbol{S}_{yy}|^M} \operatorname{etr}\left[-\boldsymbol{S}^{-1}_{yy}\boldsymbol{W}_{yy}\right] = \prod_k \frac{1}{|\boldsymbol{S}_{yy,k}|^M} \operatorname{etr}\left[-\boldsymbol{S}^{-1}_{yy,k}\boldsymbol{W}_{yy,k}\right] \tag{3.46}$$

We complete the transposition to the multi-frequency case, multiplying the probabilities at frequency f, still under the assumption that the susceptibilities do not depend on frequency.

$$\begin{aligned} &p(\{\boldsymbol{W}\}|\boldsymbol{\alpha}, \{S_{\Delta g_0}\}, \{\boldsymbol{\Sigma}_{yy}\}) \propto \\ &\propto \prod_f \frac{1}{S_{\Delta g_{0,f}}^{M_f}} \exp\left[-\frac{\boldsymbol{\eta}\boldsymbol{W}_f\boldsymbol{\eta}^H}{S_{\Delta g_{0,f}}}\right] \prod_k \frac{1}{|\boldsymbol{S}_{yy,k,f}|^{M_f}} \mathrm{etr}\left[-\boldsymbol{S}_{yy,k,f}^{-1}\boldsymbol{W}_{yy,k,f}\right] \end{aligned} \tag{3.47}$$

This expression, finally, can be employed for the application of Bayes' theorem, to infer the posterior for the $\boldsymbol{\alpha}$, $\{S_{\Delta g_0}\}$, and $\{\boldsymbol{\Sigma}_{yy}\}$:

$$\begin{aligned} &p(\boldsymbol{\alpha}, \{S_{\Delta g_0}\}, \{\boldsymbol{\Sigma}_{yy}\}|\{\boldsymbol{W}\}) \propto \\ &\propto \prod_f \frac{1}{S_{\Delta g_{0,f}}^{M_f}} \exp\left[-\frac{\boldsymbol{\eta}\boldsymbol{W}_f\boldsymbol{\eta}^H}{S_{\Delta g_{0,f}}}\right] \prod_k \frac{1}{|\boldsymbol{S}_{yy,k,f}|^{M_f}} \mathrm{etr}\left[-\boldsymbol{S}_{yy,k,f}^{-1}\boldsymbol{W}_{yy,k,f}\right] \times \\ &\times \text{prior}\left(\boldsymbol{\alpha}, \{S_{\Delta g_0}\}, \{\boldsymbol{\Sigma}_{yy}\}\right) \end{aligned} \tag{3.48}$$

Some considerations are necessary to understand the meaning of this expression better.

1. The posterior can be fully separated into two parts, which are multiplied. The first one is a full posterior for the residuals $\{S_{\Delta g_0}\}$ and the susceptibilities $\boldsymbol{\alpha}$:

$$\begin{aligned} &p(\boldsymbol{\alpha}, \{S_{\Delta g_0}\}|\{\boldsymbol{W}\}) \propto \\ &\propto \prod_f \frac{1}{S_{\Delta g_{0,f}}^{M_f}} \exp\left[-\frac{\boldsymbol{\eta}\cdot\boldsymbol{W}_f\cdot\boldsymbol{\eta}^H}{S_{\Delta g_{0,f}}}\right] \times \text{prior}\left(\boldsymbol{\alpha}, \{S_{\Delta g_0}\}\right) \end{aligned} \tag{3.49}$$

   This is the likelihood that we directly implement in our posterior estimation algorithm. As discussed, we employ Jeffrey's prior on $S_{\Delta g_0}$, which is $p(S_{\Delta g_0}) \propto 1/S_{\Delta g_0}$. We implement no prior on $\boldsymbol{\alpha}$, to be completely uninformative and unbiased. As the distribution is not simple to manipulate analytically, we revert to numerical sampling methods, as discussed later.
   We note that, if only one frequency is analyzed, with a flat prior choice, this distribution peaks at

$$S_{\Delta g_0} \to 1/P_{\Delta g_0}, \qquad \boldsymbol{\alpha} \to P_{\Delta gy} \cdot \boldsymbol{P}_{yy}^{-1}. \tag{3.50}$$

   These results are compatible with what was noted in Eqs. 3.30 and 3.34.

2. The second part in Eq. (3.48) is just the product of independent complex Wishart distributions, and it can be easily rewritten. In particular, frequencies f and blocks k completely separate into independent distributions. Employing Jeffrey's prior on each matrix, see Sec. 3.3, we end up stating that the posterior

distributions of the inverse $\mathbf{\Sigma}_{yy}$ matrices are complex-Wishart:

$$\mathbf{\Sigma}^{-1}_{yy,k,f} \sim \mathcal{CW}\left(\boldsymbol{W}^{-1}_{yy,k,f}, M_f\right) \tag{3.51}$$

As such, elements of the matrices can be drawn with the complex Wishart generator described in Appendix B.1.

3. The first observation allows us to extract samples of the posterior of the residual noise $S_{\Delta g_0}$ and the susceptibilities $\boldsymbol{\alpha}$, as we said, with numerical methods. Using the second observation, we can also extract samples of the signal CPSD matrices, with a sample generator. Putting together these two, we are able to extract samples of the distribution of the single contribution to noise of the considered time series. Given the possible cross-correlations, separation is only possible for those time series that are in different blocks in Eq. (3.45).
Assuming to extract N samples of $\boldsymbol{\alpha}$ with numerical methods, we also extract N samples of the sub-block $\mathbf{\Sigma}_{yy,k,f}$, and we calculate the contribution to noise as

$$S_{\Delta g,\mathrm{contr},k,f} = \boldsymbol{\alpha}_k \cdot \mathbf{\Sigma}_{yy,k,f} \cdot \boldsymbol{\alpha}_k^H, \tag{3.52}$$

where $\boldsymbol{\alpha}_k$ is the sub-block of $\boldsymbol{\alpha}$ corresponding to the time series included in the k-th block.

4. Equation (3.49) is also insightful for the single-frequency case. Marginalizing the likelihood with respect to $\boldsymbol{\alpha}$ in case of only one frequency, we recover the distribution of $S_{\Delta g_0}$. With Jeffrey's prior, we prove that $S_{\Delta g_0}$ follows the inverse-Gamma distribution

$$S_{\Delta g_0} \sim \mathrm{invGamma}\left(M - r/2, M\,\mathrm{Re}\mathbf{\Pi}/\mathrm{Re}\boldsymbol{P}_{yy}\right), \tag{3.53}$$

where Re indicates the real part, and $\mathrm{Re}\mathbf{\Pi}/\mathrm{Re}\boldsymbol{P}_{yy}$ is the usual Schur complement. The proof is postponed to App. B.2.3.
This result differs from that in Eq. (3.34), and we notice that the difference comes from the assumption that the susceptibilities $\boldsymbol{\alpha}$ are real. We check this expression numerically at the end of this chapter, in the test section.

5. A last observation is necessary about the applicability of these conclusions in the degenerate case, with $M < p$, i.e. with a number of periodograms which is smaller than the number of decorrelated time series. At the beginning of this section, we noted that this holds true only if $M_f \geq p$ for all f, but what happens if $M_f < p$ for some f?
At the beginning of this chapter, we noted that the complex Wishart distribution can be extended to the singular complex Wishart distribution in the degenerate case $M < p$, Eq. (3.10). Such probability density differs shows the same dependence on $\mathbf{\Sigma}$ as the complex Wishart, so that the derivation remains correct also for those frequencies for which $M < p$. Still, it is necessary that at least one among the CPSD matrices has $M \geq p$: indeed, the exponential term needs

not to be degenerate. If at least one of the terms is not degenerate, i.e. it has full rank, then the condition is satisfied[4].

There is still one caveat: the $\boldsymbol{S}_{yy}$ block in Eq. (3.44) can easily be degenerate, hence its inverse would not be complex Wishart and Eq. 3.51 could not be applied for in case the number of time series in the k-th block, r_k, satisfies $M_f > r_k$. If that is not the case, samples of the contribution of the k-th subset to $S_{\Delta g}$ can not be drawn. Still, in that case, the residuals and the susceptibilities can be estimated.

Numerical MCMC sampling

The probability distribution in Eq. (3.49) can not be easily marginalized, i.e. individual distributions of the parameters $\boldsymbol{\alpha}$ and $S_{\mathrm{res},f}$ can not be found analytically. Hence, we revert to numerical methods to generate a set of samples, distributed according to the posterior distribution. We draw samples from the marginal posteriors from this set.

The tool that we employ belongs to the family of the Markov Chain Monte Carlo (MCMC) methods, which are one of the most popular methods for distribution integration and sampling. Precisely, we implement the *parallel tempering* MCMC method, PT-MCMC [64]. PT-MCMC is a multi-layer and multi-core version of the standard Metropolis-Hastings algorithm, with variance adaptation, based on Python with MPI adaptation. It performs separate Metropolis-Hastings sampling on different layers, handled by different parallel cores. Each layer is endowed with variance adaptation for optimized sampling. Each layer samples from a different distribution, in such a way that only one layer samples from the actual target distribution, and the other one serve as support, sampling from slightly broadened versions of it. The support layers have therefore much freedom in sampling the distribution far from the high-probability peaks. From time to time, layers exchange samples among themselves, so that layers closer to the main one can access samples from the broadened ones.

This tool is widely used in GW parameter estimations, for example by the LVK and PTA collaborations, since it allows for high-dimensional sampling even with short burn-in time, and allows for efficient sampling even in multi-modal distribution cases.

We use PT-MCMC with single-component variance adaptation, using 14 parallel layers working on 14 cores, 1×10^5 burn-in samples, and 1×10^6 production samples; we subsample the production samples to reduce correlation. We note that, in the worst case, we use 26 parameters[5]. We note that we always find smooth distributions, with no bi-modalities and no relevant cross-correlations among parameters. We run the algorithm on the following set of parameters, in the following conditions:

- The residuals $S_{\Delta g_0,f}$. To enforce Jeffrey's prior, we run the MCMC algorithm

[4]The reason is that $\boldsymbol{W}$ matrices sum within the exponential term are summed, as $\boldsymbol{\eta} \cdot \left(\sum_f \boldsymbol{W}_f\right) \cdot \boldsymbol{\eta}^H$. This quadratic form is not degenerate if $\sum_f \boldsymbol{W}_f$ has full rank. It can be proved that, since the $\boldsymbol{W}_f$ are positive semidefinite, then the rank of $\sum_f \boldsymbol{W}_f$ is greater than $\max_f(\mathrm{rank}(\boldsymbol{W}_f))$. Proof can be found, for instance, at *this link*. Hence, the conclusion is that if at least one of the $\boldsymbol{W}_f$ has full rank, the condition is satisfied.

[5]The worst case corresponds to 1 Brownian level, 8 frequencies, and 17 susceptibilities, that we assume to be real.

on $\log S_{\Delta g_0,f}$, which indeed corresponds to multiplying Eq. (3.49) by the prior $1/S_{\Delta g_0,f}$.

- The real susceptibilities α_i, independently.

3.6.2 Susceptibility estimation in case of noisy signals

In Eq. (3.50), we noted that the distributions of the susceptibilities and the residuals peak at the expected theoretical value $\boldsymbol{P}_{\Delta gy} \cdot \boldsymbol{P}_{yy}^{-1}$, and it is unbiased. However, a question naturally arises in the experimental case: what happens if the signals y_i are affected by a non-negligible readout noise? What happens if we employ this model but it does not fully apply because signals are corrupted? We address this problem in the single-frequency case, analytically, proceeding by steps. Later, in Sec. 3.6.3, we validate it numerically.

We start from the expression of the CPSD matrix $\boldsymbol{\Sigma}$, in Eq. (3.26). We check what happens to the estimates of the susceptibilities $\boldsymbol{\alpha}^{(*)}$ and the residual PSD $S_{\text{res}}^{(*)}$, when the measurement is affected by a non-negligible noise. Superscripts $^{(*)}$ refer to the posterior estimates. We do not look at the estimate errors, but rather at the best estimate values, for simplicity.
We assume that the theoretical CPSD can be expressed as

$$\boldsymbol{\Sigma} = \left(\begin{array}{c|c} S_{\text{res}} + \boldsymbol{\alpha} \cdot \boldsymbol{S}_{zz} \cdot \boldsymbol{\alpha}^H & \boldsymbol{\alpha} \cdot \boldsymbol{S}_{zz} \\ \hline \boldsymbol{S}_{zz}^H \cdot \boldsymbol{\alpha}^H & \boldsymbol{S}_{zz} + \boldsymbol{S}_{nn} \end{array}\right) \tag{3.54}$$

Employing the no-noise model as in the previous sections, one would estimate the posterior values as in Eq. (3.50). Here, the presence of S_{nn} induces a bias into the estimates. Following the mathematical details presented in App. B.2.2, in particular in Eqs. B.9 and B.10, the posterior estimates of residual PSD and susceptibilities are the following[6]:

$$\begin{aligned} S_{\text{res}}^{(*)} = (\boldsymbol{A} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C}) &= S_{\text{res}} + \boldsymbol{\alpha}\boldsymbol{S}_{zz}\boldsymbol{\alpha}^H - \boldsymbol{\alpha}\boldsymbol{S}_{zz}(\boldsymbol{S}_{zz} + \boldsymbol{S}_{nn})^{-1}\boldsymbol{S}_{zz}\boldsymbol{\alpha}^H \\ &= S_{\text{res}} + \boldsymbol{\alpha}\left(\mathbb{I} - (\mathbb{I} + \boldsymbol{S}_{zz}^{-1}\boldsymbol{S}_{nn})^{-1}\right)\boldsymbol{S}_{zz}\boldsymbol{\alpha}^H \end{aligned} \tag{3.55}$$

$$\begin{aligned} \boldsymbol{\alpha}^{(*)} = \boldsymbol{B}\boldsymbol{D}^{-1} &= \boldsymbol{\alpha}\,\boldsymbol{S}_{zz}\,(\boldsymbol{S}_{zz} + \boldsymbol{S}_{nn})^{-1} = \\ &= \boldsymbol{\alpha}\,(\mathbb{I} + \boldsymbol{S}_{zz}^{-1}\boldsymbol{S}_{nn})^{-1} \end{aligned} \tag{3.56}$$

where $\mathbb{I}$ is the $p \times p$ identity matrix.

Residual PSD. The presence of noise in the signal time series affects the effectiveness of decorrelation, as in Eq. (3.55), with the following limits:

$$\text{if } |\boldsymbol{S}_{nn}| \to 0 \quad \Rightarrow \quad S_{\text{res}}^{(*)} \to S_{\text{res}} \tag{3.57}$$

$$\text{if } |\boldsymbol{S}_{nn}| \to +\infty \quad \Rightarrow \quad S_{\text{res}}^{(*)} \to S_{\text{res}} + \boldsymbol{\alpha}\boldsymbol{S}_{zz}\boldsymbol{\alpha}^H = S_{\text{tot}} \tag{3.58}$$

[6] Letters $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$ refer to the notation in App. B.2.2.

In the no-noise case, the residual PSD is fully recovered. Otherwise, the estimated residual increases, reaching the total noise S_{tot} when the noise is dominant. The latter case is equivalent to no subtraction being performed. What is remarkable is that decorrelation can not lead to oversubtraction in any case, meaning that the application of the no-noise model to noisy signals can not lead to overestimation of the contribution of signal time series to measurements. The estimated residual will be in any case higher than that in case of negligible noise.

If $|\boldsymbol{S}_{nn}| \ll |\boldsymbol{S}_{zz}|$, an additional series expansion can be performed on Eq. (3.55), resulting in

$$S_{\text{res}}^{(*)} \sim S_{\text{res}} + \boldsymbol{\alpha}\boldsymbol{S}_{zz}^{-1}\boldsymbol{S}_{nn}\boldsymbol{S}_{zz}\boldsymbol{\alpha}^{H}, \tag{3.59}$$

which gives a better grasp of the effect of noise on the estimated residuals.

In the 1D case, with only one signal y, the deviation from the correct estimate can be calculated. Let's assume that $|\rho_n|^2$ is the noise power fraction within the measured signal, $|\rho_n|^2 = S_{nn}/(S_{zz} + S_{nn})$, and that $R^2 = S_{\text{contr}}/S_{\Delta g} = \alpha^2 S_{zz}/S_{\Delta g}$ is the true contribution of y to Δg. Then, the deviation of the estimate $S_{\text{res}}^{(*)}$ from the true residual S_{res} reads

$$\frac{S_{\text{res}}^{(*)} - S_{\text{res}}}{S_{\text{res}}} = |\rho_n|^2 \left(\frac{S_{\Delta g}}{S_{\text{res}}} - 1\right) = \frac{|\rho_n|^2 R^2}{1 - R^2}. \tag{3.60}$$

This means that the fractional deviation from the true value depends on both the noise fraction $|\rho_n|^2$, and the true contribution R^2. This result is depicted in Fig. 3.5, right panel; results in the left panel are explained later.
In case z and Δg are non-correlated, i.e. $S_{\text{res}} = P_{\Delta g}$, then $S_{\text{res}}^{(*)} \to S_{\text{res}}$ independently of the noise level $|\rho_n|^2 = 1$. If the contribution is nearly negligible, then the estimate of the residual does not depend on the presence of additional noise.

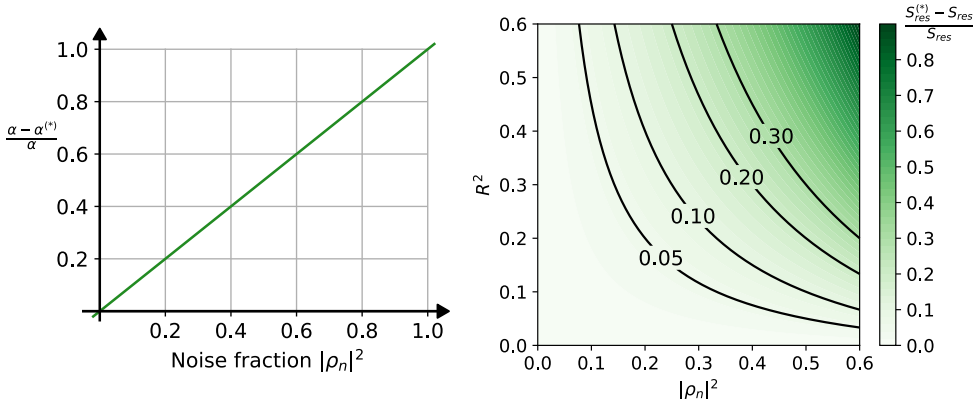


Figure 3.5: (left) Ratio between the estimated susceptibilities, depending on noise fraction $|\rho_n|^2$, the noise power fraction within the measured signal y, 1D case. The quantity $\alpha^{(*)}$ represents the posterior estimate of the susceptibility α, in case noise affects the signal time series.
(right) Ratio between the estimated residuals and the true residuals, depending on noise fraction $|\rho_n|^2$ and the true contribution of y to Δg, R^2; color-code on the right.

Susceptibilities. According to Eq. (3.56), the estimate of susceptibilities is also affected by the presence of additional noise. We note the following:

1. **1D case.** In the 1D case, Eq. (3.56) is easily approachable, as

$$\alpha^{(*)} = \alpha(1 + S_{nn}/S_{zz})^{-1} = \alpha \, \frac{S_{zz}}{S_{zz} + S_{nn}}, \tag{3.61}$$

meaning that the estimate with noise appears as shifted, by a relative bias that reads:

$$\frac{\alpha - \alpha^{(*)}}{\alpha} = \frac{S_{nn}}{S_{zz} + S_{nn}} = |\rho_n|^2, \tag{3.62}$$

where $|\rho_n|^2$ is the noise power fraction within the measured signal. Hence, the susceptibility estimate α deviates from the no-noise value α_T, with a linear dependence on the noise power fraction, as shown in Fig. 3.5 on the left panel. If the signal is contributed by a fraction of noise (let's say 20% of the total power), then it is fair to expect the susceptibility posterior to be biased, shifted from the nominal value by 20%.
We highlight the following limits:

$$\begin{aligned} \text{if } S_{nn} \to 0 \quad &\Rightarrow \quad \alpha^{(*)} \to \alpha \\ \text{if } S_{nn} \to +\infty \quad &\Rightarrow \quad \alpha^{(*)} \to 0 \end{aligned} \tag{3.63}$$

2. **No correlation case.** If the signals do not have any correlation among them, i.e., if the off-diagonal elements of the matrix $\boldsymbol{S}_{zz}$ are zero, elements in Eq. (3.56) turn out to be independent:

$$\alpha_i^{(*)} = \alpha_i \, (1 + S_{n_i n_i}/S_{z_i z_i})^{-1}, \tag{3.64}$$

hence the problem reduces to the previous one, in the 1D case.

3. **2D case with correlation.** In the most general case, signals z might cross-correlate one to another, being affected by additional readout noises, which do not correlate. For simplicity, let's suppose that $\boldsymbol{S}_{zz}$ is a 2D matrix, built with the signals v and w, showing a cross-correlation ρ, with additional noises:

$$\boldsymbol{S}_{zz} = \begin{pmatrix} S_v & \rho\sqrt{S_v S_w} \\ \rho^*\sqrt{S_v S_w} & S_w \end{pmatrix} \qquad \boldsymbol{S}_{nn} = \begin{pmatrix} N_v & 0 \\ 0 & N_w \end{pmatrix} \tag{3.65}$$

In this case, the estimate of susceptibilities, as in Eq. (3.56), is a bit harder to read; we provide the expression for $\alpha_v^{(*)}$, which is

$$\alpha_v^{(*)} = \frac{\alpha_v S_v (N_w + S_w(1 - |\rho|^2) + \alpha_w N_v \rho^* \sqrt{S_v S_w}}{(N_v + S_v)(N_w + S_w) - S_v S_w |\rho|^2} \tag{3.66}$$

An analogous expression holds for $\alpha_w^{(*)}$. We note a few cases:

3.1. $N_v \to 0$.

If the signal v is not affected by any readout noise, then the estimate of its susceptibility is unbiased,

$$\alpha_v^{(*)} \to \alpha_v \tag{3.67}$$

3.2. $N_w \to 0$.
If the signal w is not affected by any readout noise, then its PSD affects the estimate of the susceptibility of the signal v, as

$$\alpha_v^{(*)} \to \alpha_v \frac{1}{1 + \frac{N_v}{S_v(1-|\rho|^2)}} + \alpha_w \frac{N_v \rho^* \sqrt{S_v S_w}}{S_w(N_v + S_v(1-|\rho|^2))} \tag{3.68}$$

3.3. $|\rho| \to 0$.
If the correlation ρ is zero, we find case #2.
However, in this case one can expand in series Eq. (3.66) with respect to ρ, looking at the case of low-correlation among signals. This yields the correction given by the correlation, which is:

$$\alpha_v^{(*)} \to \alpha_v \frac{S_v}{S_v + S_n} + \alpha_w \frac{N_v \sqrt{S_v S_w}}{(N_v + S_v)(N_w + S_w)} \rho^* + \mathcal{O}[|\rho|^2] \tag{3.69}$$

3.4. $N_w \to 0, |\rho| \to 1$.
In case no noise is present on N_w, but the correlation between v and w is the highest possible, $|\rho| \to 1$, then the susceptibility on v can be fully recovered, even if readout noise is affecting the signal v. In formulas,

$$\alpha_v^{(*)} \to \alpha_w \sqrt{S_v / S_w} \tag{3.70}$$

Note that in this case, $\alpha_w^{(*)} \to \alpha_w$ as $N_w \to 0$, hence knowing S_w is enough to recover both the susceptibilities. However, one would need to know in advance that $|\rho| = 1$, as this information is otherwise hidden by noise N_v.

3.6.3 Numerical tests

In order to consolidate the methods presented in this chapter, and in particular the multi-frequency decorrelation, we perform some numerical tests. The purpose of the tests is

1. to check the compatibility between the multi-frequency decorrelation method and the single-frequency one, in case only one frequency is taken into account;

2. to check that, in a controlled environment with synthetic time series and no additional readout noise, the residual noise and the susceptibilities are accurately recovered;

3. to check that, with synthetic time series, adding a low percentage of synthetic readout noise does not prevent using the decorrelation method; this test is also used to understand how large the signal readout noise should be to spoil the parameter estimation.

Test #1: single-frequency case. First, we test the compatibility of the single-frequency and multi-frequency methods in the case of only one frequency being analyzed. We simulate two cross-correlating time series, and decorrelate one with respect to the other. In Fig 3.6, we show the histogram of the residual noise evaluated with the MCMC model in Eq. (3.49), and its compatibility with the expected analytical distribution. We check two cases:

- A first case, in which $\boldsymbol{\alpha}$ is complex, and we hypothesize that the posterior distribution of the residual follows Eq. (3.34). That case is represented on the left in Fig. 3.6.
- A second case, in which $\boldsymbol{\alpha}$ is real, and we hypothesize that the posterior distribution of the residual follows Eq. (3.53). That case is represented on the right in Fig. 3.6.

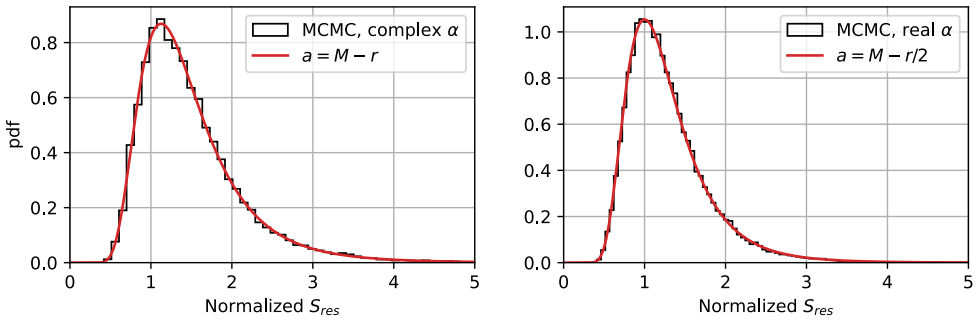


Figure 3.6: Comparison between the MCMC numerical multi-frequency model and the analytical single-frequency model, with synthetic simulated time series (9 periodograms in this simulation). The red line is the residual noise, normalized to the nominal value, of the residual after decorrelation in the single-frequency case, with Jeffrey's prior. The left and right panel refer, respectively, to the complex and the real case.

Test #2: multi-frequency case with no readout noise. To test the multi-frequency method, we simulate synthetic time series, lasting as long as run Feb17B, with the tools provided by LTPDA [65]. We build the time series as described in App. B.3. We simulate a main measurement, shaped as the LPF noise, and two signals, respectively following power-law shapes $1/f$ and $1/f^2$. The signals cross-correlate with one another, and contribute additively to the main series through two susceptibilities. We simulate two cases:

1. A high-corruption case, meaning that the signals y_i contribute up to 75% of the total measured main signal power. In this case, we aim at recovering the residual, hidden under the corruption of the two signals. The used signals are represented in Fig. 3.8, on the left.
2. A low-corruption case, with signals contributing at most to the 5-10% of the total main signal power. This is the most LPF-resembling case, in which we aim at recovering the susceptibilities $\boldsymbol{\alpha}$, even if the signal is low-power.

We find good results in both cases. In Fig. 3.7, we provide the results of the MCMC analysis in the first case. We use the posterior in Eq. (3.49), with Jeffrey's prior on the residuals. Resembling the LPF case, we only use real susceptibilities, including 9 frequencies from 18 µHz to 1.3 mHz.

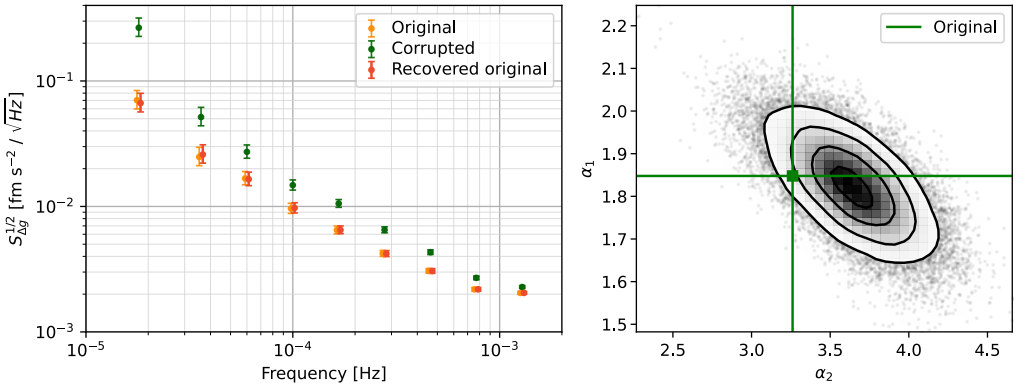


Figure 3.7: Numerical test of the decorrelation method, as described in the text. The original signal is corrupted with additional noises a and b, coupling with susceptibilities $\boldsymbol{\alpha}$.
(left) The original signal (orange) is corrupted (green) with a and b, as in Fig. 3.8 on the left. The red points represent the recovered signal after decorrelation.
(right) Recovered susceptibilities α_1 and α_2. The green lines represent the 'true' value. We note that the distribution peaks at $\boldsymbol{\alpha} = \boldsymbol{P}_{gy} \cdot \boldsymbol{P}_{yy}^{-1}$, as expected.

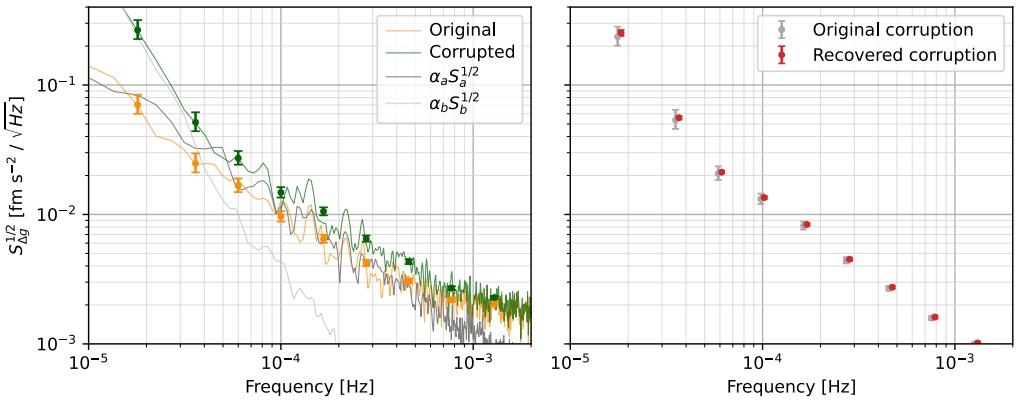


Figure 3.8: (left) Contributions from the single terms a and b to the total noise $S_{\Delta g}$.
(right) Reconstruction of the noise contribution from a and b: original (grey) and estimated (red). As noted in the text, we underline that this is not a true posterior, but rather a rough estimate used to distinguish the individual contributions.

Test #3: multi-frequency case with readout noise. In addition to test #2, we add spurious readout noise to the signals y_i, and apply the multi-frequency decorrelation method. We repeat the tests with increasing levels of noise, built as described in App. B.3. We note that the model does not include the presence of any readout noise, hence we do not expect it to be precise with high levels of noise. The reason for such tests is that in the LPF case, even though we expect the signals to

be affected by negligible noise at sub-mHz frequencies, we can not exclude a (small) percentage of noise.

We find that, in case the readout noise of the signals contributes up to 15% of the total signal power, the residual noise $S_{\Delta g_0}$ is well-recovered.

4 Low-frequency band: performance and noise nature

PSD analysis, time evolution and stability, noise nature

In this chapter, we present a preliminary analysis of the LPF acceleration performances. We consider a set of 14 long runs, lasting longer than 2.75 days, and analyze the $\Delta g(t)$ time series. We present observations about measurements, noise performance, and the noise PSD stability during the LPF mission.

From previous examinations, it is known that the sub-mHz Δg noise may be effectively split into branches: the high-frequency noise, above mHz frequencies, is confidently ascribed to interferometer noise. The sub-mHz band may also be divided into two branches, i.e., a high-frequency white-noise plateau, and a low-frequency branch scaling as $f^{-2.0\pm0.2}$ [2, 3]. First, we separate these two branches. Then, we provide analyses of the nature of the detected noise, and its properties, such as its stability, its evolution in time, and its Gaussianity.

The analyses presented here are largely based on what is presented in [2, 3], providing extended and refined analyses. The results of these analyses will be presented in [5].

4.1 PSD analysis: Brownian noise and excess

4.1.1 Run selection

We start the analysis by computing the PSDs. We choose a set of 14 noise-only runs, marked in orange in Table A.1, distributed all over the LPF mission. The first considered run, Mar16E, is the only one performed with `URLA` actuation settings, and it is performed about 40 days after venting to space. Run Apr17A is the only one performed with DRS `DURLA` settings, and it is used for comparison between DRS and LTP runs. The remaining runs were performed with LTP `UURLA` settings in stationary experimental conditions, with no applied external stimuli. These 14 runs last at least 2.75 d (Jul16C), allowing to analyze frequencies down to 18 µHz with at least 1 periodogram at the lowest frequency. Among the analyzed runs, we often represent run Feb17B as a reference, as in [2, 3]. Indeed, run Feb17B is one of the most representative, as it is one of the best performances achieved during the mission[1].

[1] As we will comment later, Feb17B is one of the best performances achieved, under the point of view of low-frequency noise level, stability, and Brownian noise level.

For each run, we compute PSDs and ASDs of the Δg time series, defined in Eq. (2.6), and corrected for inertial effects, cross-talks, and deglitched. In Figure 4.1, blue points represent the measured ASD in run Feb17B. PSDs of other runs in the considered set are shown in Figs. 4.2, 4.3, and 4.4, focusing on the frequency band [18 µHz–10 mHz].

4.1.2 Brownian plateau identification

In each of these noise runs, it is evident that $S_{\Delta g}$ is generally characterized by three branches:

- A high-frequency branch, not interesting to our purposes, showing up above 10 mHz, dominated by the interferometer readout noise [2, 13, 33]. Especially at the beginning of the mission, this feature is not visible in the figures in this chapter, as the white-noise plateau is often hiding it.
- A low-frequency (< 0.1 mHz) branch, characterized by a frequency dependence $\propto f^{-2}$ as shown in [3].
- A central white-noise plateau, dominating noise at frequencies $\sim$ mHz, with a magnitude that decreases in time.

First, we separate the white-noise plateau, confidently ascribed to Brownian noise (analyzed later, Chapter 5), from the low-frequency excess. Precise statistical inference is necessary not to estimate the Brownian magnitude, which is rather clear and robust, but rather the excess noise, which is difficult to identify in presence of the strong white noise. We estimate the excess over white noise up to $f = 2.1$ mHz.The estimates at higher frequencies are unreliable, especially in case of high white noise, and not needed. We model the system assuming that:

1. At each frequency bin, the PSD $S_{\Delta g}(f)$ is contributed by the frequency-independent white noise S_B, and the frequency-dependent excess $S_e(f)$.
2. The frequency bin with minimum PSD, between 18 µHz and 10 mHz, is only contributed by the white, Brownian noise S_B.
3. The measured noise at each frequency, $S_{\Delta g}(f)$, follows the complex-Wishart distribution[2], as described in Chapter 3, with a theoretical PSD $S_B + S_e(f)$.
4. We use Bayesian statistics to estimate S_B and $S_e(f)$. We assume a flat prior on the logarithms of the inferred parameters S_B and $S_e(f)$, following Jeffrey's non-informative prior.

In formulas, the model reads

$$\begin{aligned} p\left(S_B, \{S_e\} | \{S_{\Delta g}\}\right) = \prod_{\substack{i\in[1,10] \\ i\neq i_{\min}}} \Big[p\left(S_B + S_e(f_i) | S_{\Delta g}(f_i)\right) \times \mathrm{pr}(S_e(f_i)) \Big] \times \\ \times\, p\left(S_B | S_{\Delta g}(f_{i_{\min}})\right) \times \mathrm{pr}(S_B) \end{aligned} \tag{4.1}$$

[2] As the one-dimension PSD is real, it can either be modeled as complex or real Wishart distributions. Both lead to the same Gamma distribution, as in Eq. (3.11).

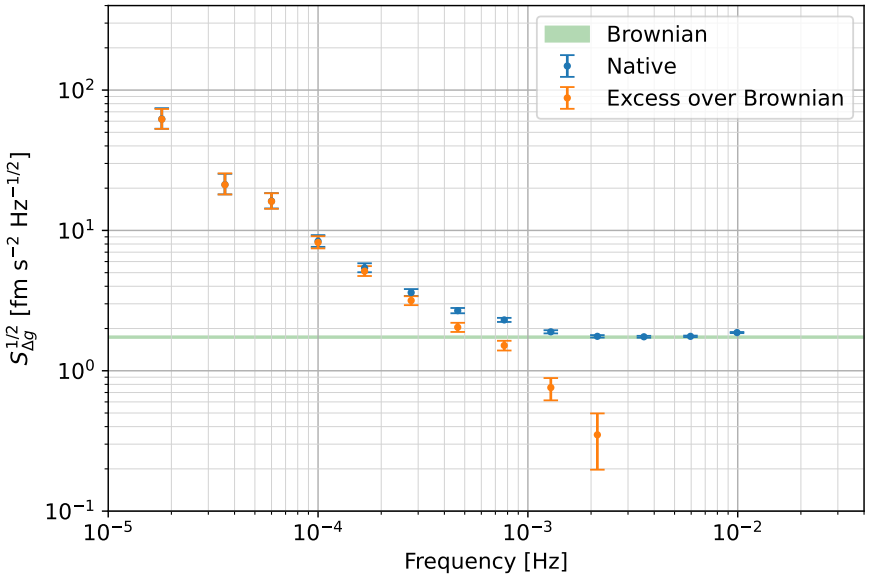


Figure 4.1: Identification of the f^{-2} and white-noise branches in $S^{1/2}_{\Delta g}$, run Feb17B. Blue points represent the measured native ASD. The green line and the orange points represent respectively the estimate of the white-noise plateau noise and the excess noise above it.

where p on the right-hand side is $p(S|\Pi) \sim \text{invΓ}(M-1, M\Pi)$, M the number of periodograms, and the prior $\text{pr}(S) \propto 1/S$.

We use the probability density in Eq. (4.1) to infer the parameters S_B and $S_e(f)$, for each frequency. As usual, we employ an MCMC method with a parallel-tempering algorithm (PT-MCMC). Results for run Feb17B, as an example, are shown in Fig. 4.1: the green line is the 1σ estimate (16%–84% quantiles) of the Brownian level, and the orange points are the estimates of the excess above it.

At very low frequencies, it is obvious that the difference between native and excess is minimal; however, it is significant at $\sim$ mHz frequencies. We note that, given the model in Eq. (4.1), and given that the number of averaged periodograms at mHz frequencies is some hundreds, the estimate of Brownian is basically given by the second line in Eq. (4.1), regardless of $S_e(f)$. Therefore, this MCMC method is not essential to estimate S_B, but it is to estimate $S_e(f)$, the excess noise over Brownian.

In Figs. 4.2, 4.3, and 4.4, we provide the results of Brownian and excess estimation for the entire set of the 14 runs considered.

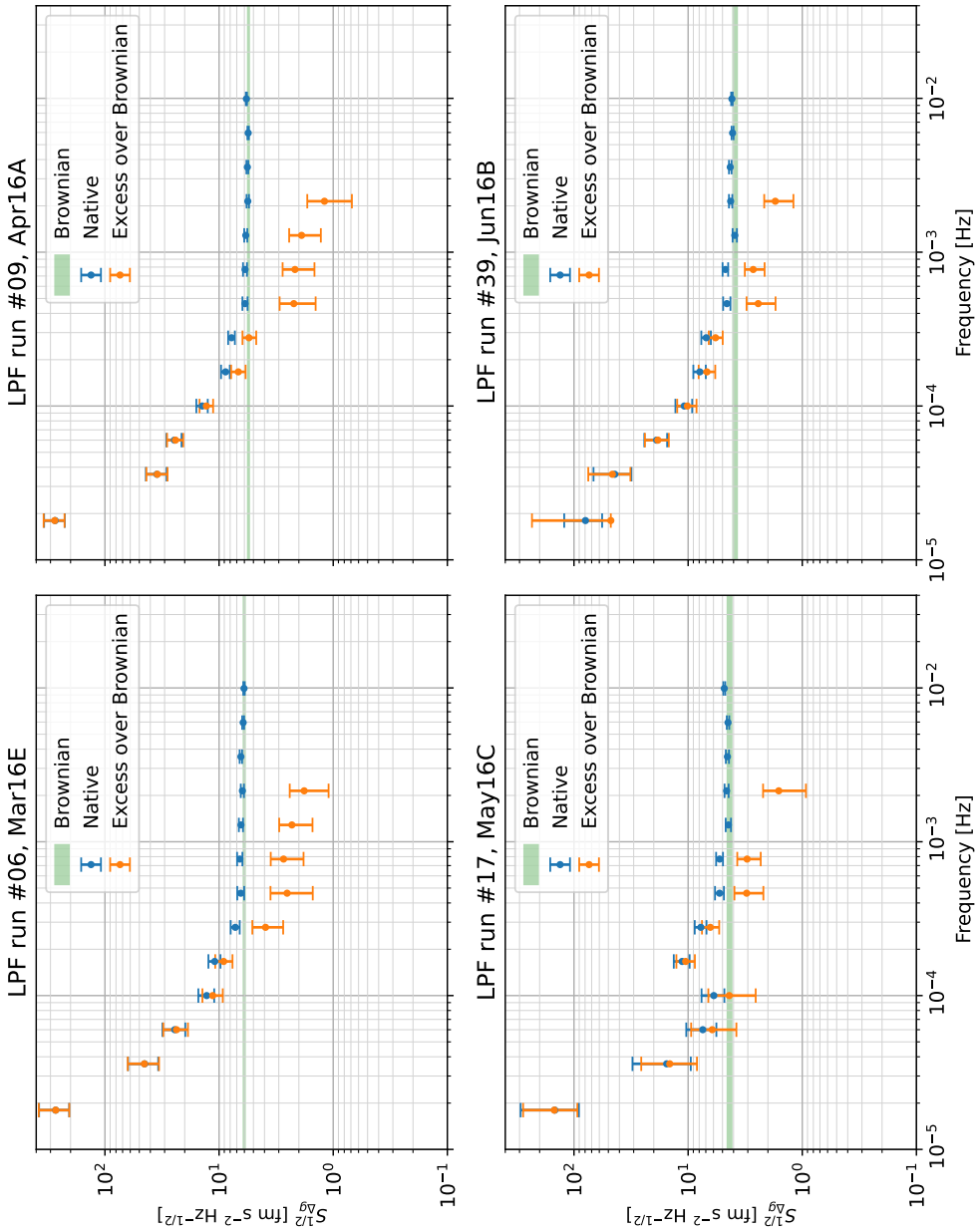


Figure 4.2: Identification of the f^{-2} and white-noise branches in $S_{\Delta g}^{1/2}$, runs Mar16E (URLA), Apr16A, May16C, Jun16B. Blue points represent the measured native ASD of Δg. The green line and the orange points represent respectively the estimate of the white-noise plateau noise and the excess noise above it.

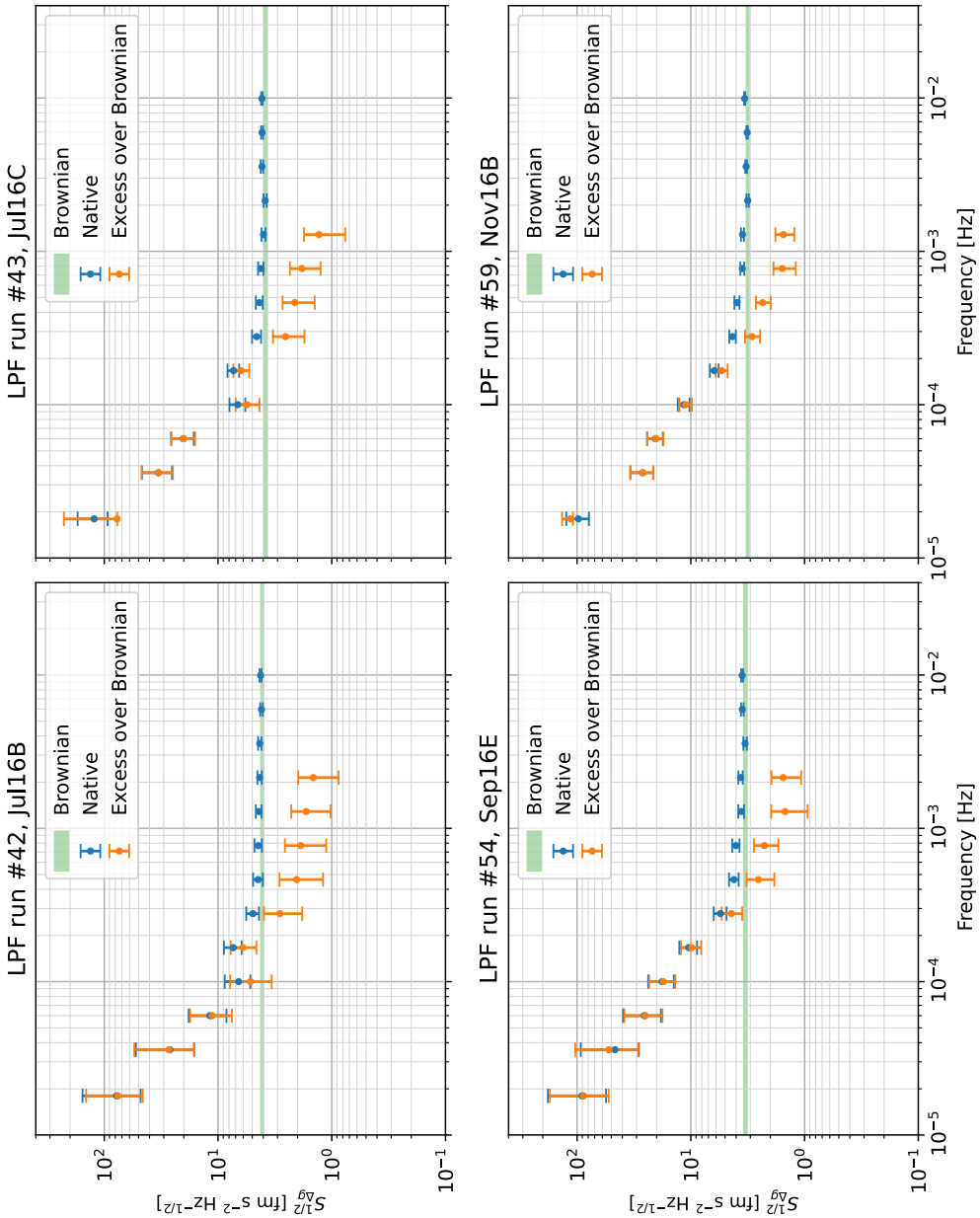


Figure 4.3: Identification of the f^{-2} and white-noise branches in $S_{\Delta g}^{1/2}$, runs Jul16B, Jul16C, Sep16E, Nov16B. Blue points represent the measured native ASD of Δg. The green line and the orange points represent respectively the estimate of the white-noise plateau noise and the excess noise above it.

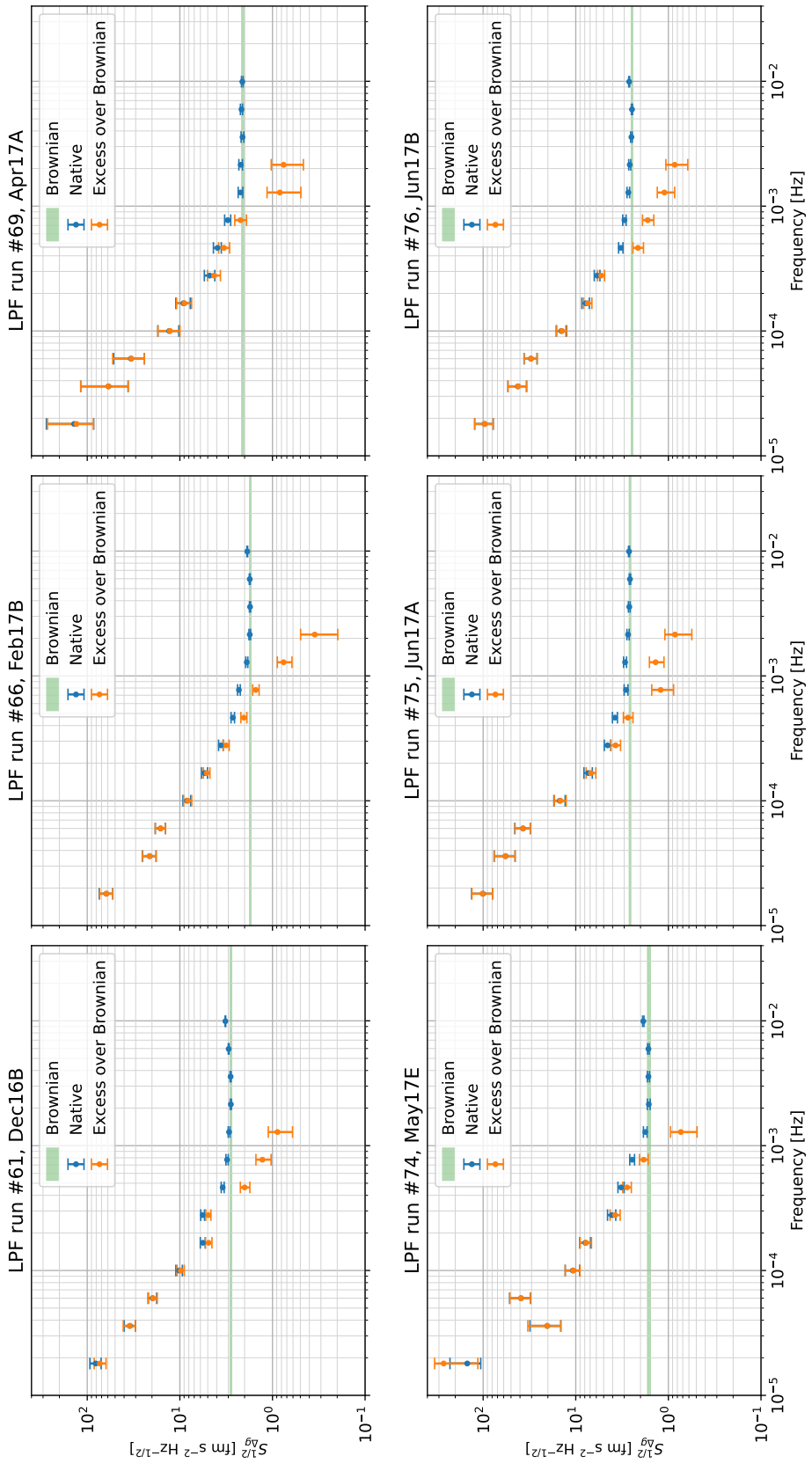


Figure 4.4: Identification of the f^{-2} and white-noise branches in $S^{1/2}_{\Delta g}$, runs Dec16B, Feb17B, Apr17A (DRS), May17E, Jun17A, Jun17B. Blue points represent the measured native ASD of Δg. The green line and the orange points represent respectively the estimate of the white-noise plateau noise and the excess noise above it.

4.2 Properties of the excess noise

We divide the analyses about the excess noise S_e and the white-noise level, starting with the former. The evolution of the white-noise level will be discussed thoroughly in Chapter 5.

Looking at the PSDs of the excess noise in Figs. 4.2 to 4.4, two more questions come up to mind about the nature of the measured noise. The first is if there is any chance that the detected excess is common among all runs, and the detected fluctuations are just caused by the intrinsic Wishart statistics. The second question is a consequence of the first one: if noise is not common to all runs, what is its stability over the mission duration?

4.2.1 Fit to a common excess noise

To answer the first question, we first need to fit observation to a common-noise model. We adapt the single-run model in Eq. (4.1), multiplying single-run probabilities to build a collective posterior. Brownian noise $S_{B,k}$ depends on the run index k, and the excess $S_e(f)$ only depends on frequency, being common to all runs. The model becomes

$$p\left(\{S_B\},\{S_e\}|\{S_{\Delta g}\}\right) = \prod_{k\in\text{runs}} \prod_{\substack{i\in[1,10]\\ i\neq i_{\min}}} \Big[p\left(S_{B,k}+S_e(f_i)|S_{\Delta g,k}(f_i)\right) \times \text{pr}(S_e(f_i)) \times \\ \times p\left(S_{B,k}|S_{\Delta g,k}(f_{i_{\min}})\right) \times \text{pr}(S_{B,k}) \Big] \tag{4.2}$$

We choose non-informative priors on the PSD parameters, and run the MCMC on the subset of 13 runs in Table A.2. We do not include run Apr16A (DRS) in this analysis, since it was performed in too different experimental conditions. The results are provided in Fig. 4.5, black points. In the same plot, we show the single-run excess PSDs, as cumulative quantiles of the single-run estimates evaluated in the previous section.

As usual in Bayesian analysis, the likelihood maximization done for fitting is not informative about the suitability of the model, it is useful just to find the best parameter set. To assess if the model is predictive, we perform a posterior predictive check [66]. The check allows to test how likely it is to find the observed experimental data, given the best-fit parameters.

First, we find the set of parameter values $\{S_{B,\text{best}}\}$, $\{S_{e,\text{best}}\}$, from the MCMC chain, maximizing a likelihood function. The log-likelihood Λ is a measure of the discrepancy between observations and the model. We compute the log-likelihood of the observed data as

$$\Lambda = \sum_{\substack{k\in\text{runs}\\ f\in\text{freqs}}} \log p\left(S_{\Delta g,k}(f)|\{S_{B,\text{best}}\},\{S_{e,\text{best}}\}\right) \tag{4.3}$$

where $p(\Pi|S)$ is the one-dimensional Wishart distribution, defining the distribution

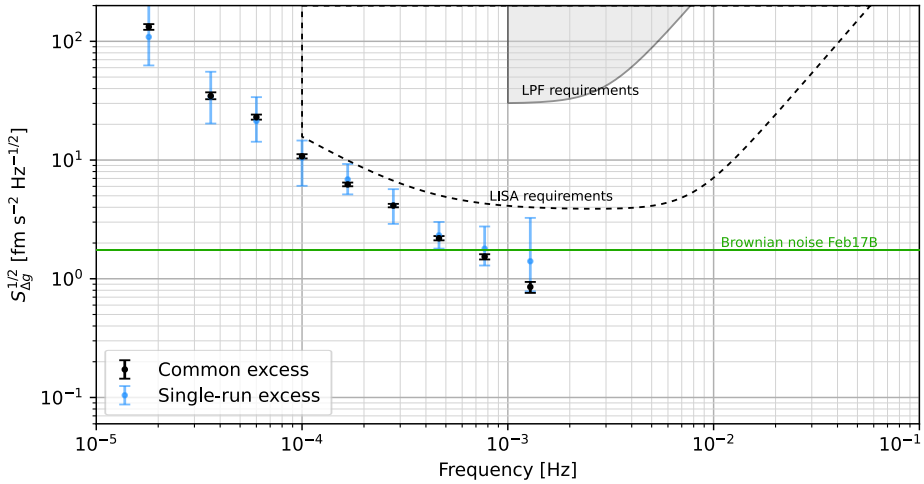


Figure 4.5: Common excess noise over Brownian, in ordinary runs in Table A.2. The common noise (black points), is evaluated with the model in Eq. (4.2). Blue points represent the single-run excess PSDs, evaluated as cumulative quantiles of the single-run estimates evaluated with the model in Eq. (4.1). For comparison, we also show the LPF and LISA requirements, and the Brownian noise level in run Feb17B.

of experimental data (see Eq. (3.11)). We discard the first frequency from the frequency set, for reasons underlined in Sec. 3.5. Then, we create a set of simulated periodograms: for a set of parameters in the posterior MCMC chains, we generate simulated periodograms from the appropriate Wishart distribution, using the same number of averaging periodograms as in single runs. We repeat the procedure for a number of parameter sets and, for each of them, we then calculate the log-likelihood Λ. The probability that Λ is less than the value calculated with real data, within this set of simulated data, is a measure of the predictiveness of the model.

We find a few results:

1. The model is not predictive of the entire run set, with $p < 0.1\%$. A common noise model of all the LPF runs does not describe measurements.

2. The model is more predictive ($p > 5\%$) if subsets of runs are chosen, instead of the entire set of 13 runs. On the one hand, this fact is compatible with noise being stable over short time-scales; on the other hand, it might just be a hint that the posterior check is not as powerful when small sets of runs are considered. In any case, we find that the test is predictive if runs are minimally sub-divided as [Mar16E, Apr16A, May16C, Jun16B], [Jul16B, Jul16C, Sep16E, Nov16B], [Dec16B, Feb17B], [May17E, Jun17A, Jun17B].

3. The common-excess model is applicable also if the excess is not strictly stable in time, but its amplitude fluctuates. In case of random instabilities of the excess noise, fluctuations as $S_e \to S_e(1+x)$, observations would fit a quasi-common excess if x is normally distributed with $\sigma_x = 25\%$.

4. However, the test is not conclusive regarding the existence of either epochs or fluctuations, or both of them.

4.2.2 Stability of the excess noise

The results from the common-noise tests hint that some fluctuations might be present in the "true value" of the low-frequency PSDs. We proceed in analyzing the magnitude of such fluctuations. To do it, we first define a figure of merit, summarizing the overall behavior of the excess noise, and capturing the most important properties. We take advantage of the fact that the excess noise scales as f^{-2} (see [2, 3]), and fit the measured PSD in each run to the model

$$S_{\Delta g}(f) = \widetilde{S}_{\Delta g} \left(\frac{1\,\mathrm{mHz}}{f} \right)^2 + S_B, \tag{4.4}$$

where S_B represents the Brownian level, and $\widetilde{S}_{\Delta g}$ is the newly-defined figure of merit. We will focus on $\widetilde{S}^{1/2}_{\Delta g}$, which represents the effective ASD of a f^{-1} fit to observations, evaluated at 1 mHz.

The fit is performed as usual, with Bayesian estimation of $\widetilde{S}_{\Delta g}$ and S_B, based on the Wishart statistics governing the PSD distributions. For each segment, we run an MCMC algorithm to infer the distribution $p(\widetilde{S}_{\Delta g}, S_B | S_{\Delta g})$ as follows

$$\begin{cases} p(\widetilde{S}_{\Delta g}, S_B | S_{\Delta g}) = \prod\limits_{f \in \mathrm{freqs}} \left[p(S^* | S_{\Delta g}) \right] \times \mathrm{pr}(\widetilde{S}_{\Delta g}) \times \mathrm{pr}(S_B) \\ S^* = \widetilde{S}_{\Delta g} \left(1\,\mathrm{mHz}/f \right)^2 + S_B \end{cases} \tag{4.5}$$

As usual, we choose an non-informative Jeffrey's prior on $\widetilde{S}_{\Delta g}$. The prior on S_B, instead, is the posterior which we found in the single-run Brownian evaluation (see Sec. 4.1). In the right-hand side of the first equation, p is the posterior 1-D Wishart distribution $p(S|\Pi) \sim \mathrm{inv}\Gamma(M-1, M\Pi)$.
We perform this fit on the set of 14 long-lasting runs, which we show in Fig. 4.6.

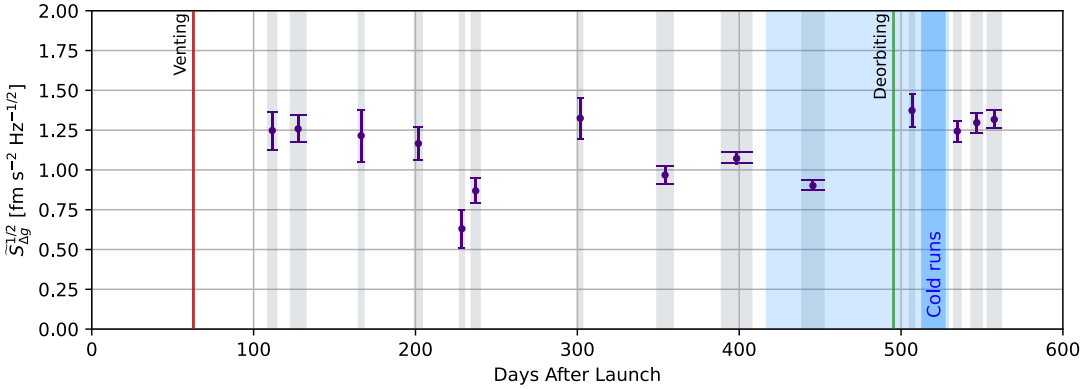


Figure 4.6: Slope of the f^{-2} sub-mHz branch, $\widetilde{S}^{1/2}_{\Delta g}$ as described in text, during the set of 14 long-lasting noise-only runs, marked in orange in Table A.1. The width of the points, also highlighted with grey spans, marks the run duration. The blue spans, the red line, and the green line, highlight respectively lower nominal temperature, the venting-to-space time, and the spacecraft deorbiting.

However, we notice that the duration of runs is very different and this averaged quantity might be hiding some internal fluctuations within the single runs. For this reason, we choose a more consistent approach, partitioning all runs into non-overlapping segments, lasting 2.75 d, which is the duration of the shortest run in our set[3]. The result of the fitting procedure performed on the segments is shown in Fig. 4.7.

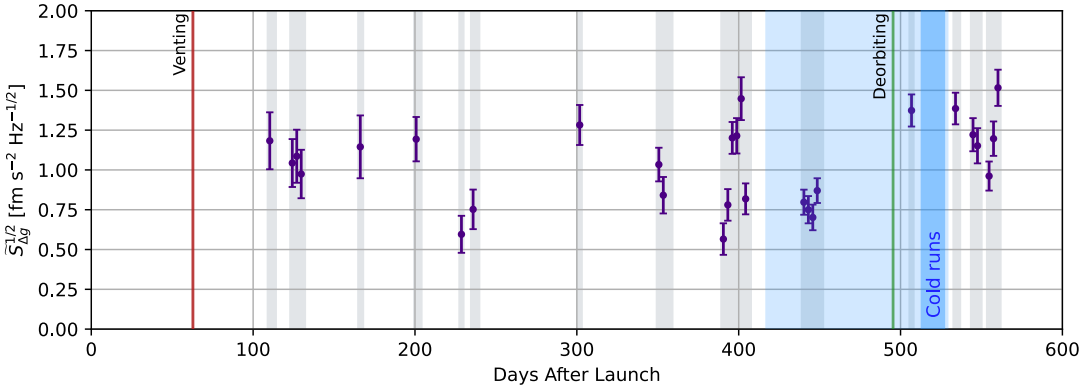


Figure 4.7: Slope of the f^{-2} sub-mHz branch, $\widetilde{S}^{1/2}_{\Delta g}$ as described in text, during the set of 14 long-lasting noise-only runs, marked in orange in Table A.1. Each run has been partitioned into 2.75 d non-overlapping segments, lasting as long as the shortest ordinary runs Jul16C. The blue spans, the red line, and the green line, highlight respectively lower nominal temperature, the venting-to-space time, and the spacecraft deorbiting.

Among the analyzed 28 segments, fluctuations are well visible. In particular, data show deviations as large as $\sim 25\%$; such a deviation is not compatible with random fluctuations due to Wishart statistics, confirming our hypothesis that the noise level was not stationary across the entire mission. Instead, we conclude that it was stationary within a remarkable 25%.

The sub-division into shorter segments gives some more details, which we underline:

1. The fluctuation amplitude observed during the entire mission is also observed during a single run, Dec16B, which incidentally is the longest run. Within that run, the excess noise level varies between the lowest observed and the second-highest observed. This shows that the low noises measured in runs Jul16B and Jul16C might just be fluctuations, rather than periods with low noise.

2. Run Feb17B, often taken as a reference run for its noise level, for its stability, and for its PSD smoothness, is the one with the smallest scatter among segments. Compatible, however, with some segments of the previous run Dec16B.

3. After cold runs, a higher noise level might be noticed, though still compatible with some points of run Dec16B. The question arises whether the noise level significantly increased, or if the variation could just be a random fluctuation. We perform a classical permutation test: given two sets of data, we evaluate the

[3]The shortest run in the subset is Jul16C, lasting about 2.75 d. However, we include run Apr16A as an exception. Even though it lasts less than 2.75 d, it is meaningful to include it in analyses, to estimate the low-frequency excess in DRS runs.

difference between the mean values of the two sets; the difference between Feb17B and May17-Jun17 is $0.45\,\mathrm{fm\,s^{-2}\,Hz^{-1/2}}$. We repeat the same operation on other data groups: we make random permutations of all 27 partitions (excluding the DRS run), and for each subset of 10 adjacent values we evaluate the difference between the mean value of a first subset of 4 data (as in Feb17B) and the mean value of a second subset of 6 data (as in May17-Jun17), and pick the maximum observed step. Then, we assess the probability of finding a difference higher than the observed one. We find that, though the set of Feb17B values is significantly lower than the ones in May17-Jun17, the observed step is still compatible with random fluctuations, as the probability of finding a maximum step greater than the observed one is 9%. We conclude that the low noise found in Feb17B could just be the outcome of a fortunate coincidence.

4.2.3 Low-frequency data Gaussianity

To conclude the analysis of noise stability, we analyze data Gaussianity. A simple explanation of the observation of a noise level fluctuating more than foreseen by the Wishart statistics could simply be that noise is not Gaussian, hence not following exactly the Wishart statistics. We run a simple test to assess low-frequency data Gaussianity.

First, we choose to test frequencies up to the 9th, 2.1 mHz, and the 13 long-lasting ordinary runs. For each run, and each frequency, we evaluate the 50%-overlapping periodograms. We normalize the real and imaginary parts of the periodograms to their mean and standard deviations within each sample. It is known that, if data are Gaussian, the real and imaginary parts of the periodogram should follow the same zero-mean Gaussian distribution, with the same standard deviation. Building standardized periodograms as we do, we expect them to follow a normal distribution with unit variance.

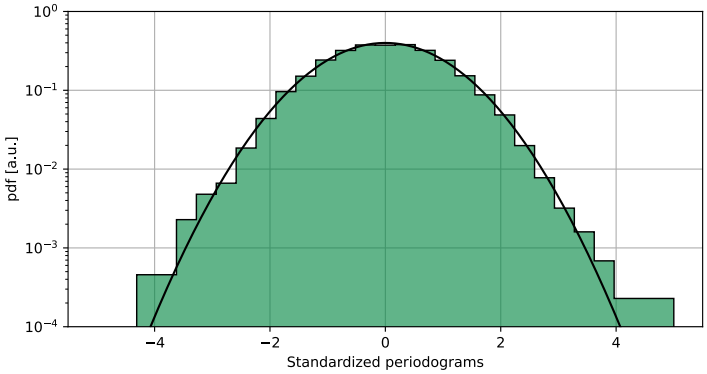


Figure 4.8: Histogram of the 12696 samples used for the Gaussianity test, as described in text, including 13 ordinary LTP runs and 9 frequencies for each of them, 18 µHz to 2.1 mHz. The black line is a normal zero-mean, unit-variance distribution, to be used as reference.

Putting together all samples from different runs and frequencies, we build a set of 12696 samples. Such a set is shown in Fig. 4.8; for comparison, we show the zero-mean unit-variance normal distribution, which shows good agreement with data. No outliers

above 5σ are detected.
We also perform an Anderson-Darling test for Gaussianity, which claims compatibility between our observations and the normal distribution, with $p > 10\%$. We run this test on the subset of the first four frequencies, up to 0.1 mHz, in which 788 samples are available. For the sake of completeness, we run the Anderson-Darling test to all the possible frequency subsets: it succeeds every time, with $p > 5\%$; in only one outlier case, the result is $p \sim 1\%$, but still not significant to exclude Gaussianity. The overall response is positive, allowing us to conclude that excess noise is compatible with Gaussian data.

4.3 Nature of the excess noise

Before moving to the analysis of white-noise evolution, some more questions arise: what is the nature of the detected excess? May it be interferometer readout noise or is it generated by real forces acting on the TMs? If it is caused by forces, does it have a torque counterpart? We try to give an answer to these questions, with two analyses. Finally, we make an observation about instrument distortion.

4.3.1 Interferometer low-frequency readout noise

First, we investigate if the excess noise might be originating from noise in the interferometric readout, or if it is due to real forces. According to the definition of Δg, Eq. (2.6), any noise in the interferometer readout, $n_{\mathrm{OMS}}(t)$ would enter the measurement, inducing a noise

$$\Delta g_n(t) = \ddot{n}_{\mathrm{OMS}}(t) + \omega_2^2 n_{\mathrm{OMS}}(t) \tag{4.6}$$

Here we compare measurements of $S_{\Delta g}$ with estimates of the interferometer readout noise, to assess its impact.

We identify two methods to assess the contribution of OMS readout noise; the lead to the same conclusion, i.e., that the detected noise is caused by real forces.

1. In April 2017, a 2.4 d run was performed with both TMs grabbed by blocking mechanisms, and no electrostatic actuation. During these runs, TM position was not controlled by the interferometer output. Hence, the interferometric measurement $\Delta x_{12,\mathrm{OMS}}$ can be used for estimation of the open-loop interferometer noise concurring to Δg. Assuming that the interferometer noise is $S_{\Delta g_n}(f)$, according to Eq. (2.6), during these runs it is given by

$$S_{\Delta g_n}(f) = S_{\Delta x_{12,\mathrm{OMS}}}(f) \left(4\pi^2 f^2 + |\omega_2^2|\right)^2 \tag{4.7}$$

   This measurement is however an upper limit of the interferometer noise, for two reasons noted in [33]:

   - It includes mechanical movements due to the rigid contact with the GRS mechanisms, likely related to thermal fluctuations.

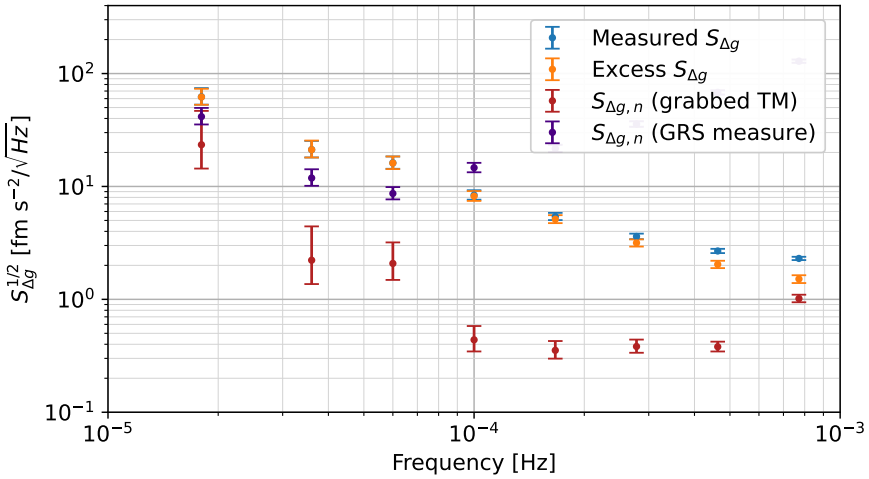


Figure 4.9: Measured noise (blue) and excess over Brownian (orange) during run Feb17B, compared to the estimates of $S_{\Delta g_n}$ (red, with grabbed TMs, and blue, with capacitive measurements). $S_{\Delta g_n}$, defined in Eq. (4.7), represents an upper limit to the interferometer readout noise.

- It was performed in low-contrast conditions, hence with a large interferometer noise, because of non-optimal surface alignment.

Figure 4.9 shows the measured noise and excess during run Feb17B, along with the estimate of $S_{\Delta g_n}$ in April 2017.

2. The second measurement is completely independent. We evaluate the out-of-loop $\Delta g(t)$ independently, with the OMS and GRS measurements. The equation governing the system is

$$\begin{cases} \Delta g_{\mathrm{OMS}}(t) &= \Delta\ddot{x}_{\mathrm{OMS}}(t) - g_{\mathrm{ext}}(t) \\ \Delta g_{\mathrm{GRS}}(t) &= \Delta\ddot{x}_{\mathrm{GRS}}(t) - g_{\mathrm{ext}}(t) \end{cases} \tag{4.8}$$

where $g_{\mathrm{ext}}(t)$ are the external forces, comprehending both actuation forces, and other truly external forces, which are common to the interferometric and capacitive readouts. It results that the difference only contains the readout noises, representing an upper limit to the OMS noise.

$$\Delta g_{\mathrm{GRS}}(t) - \Delta g_{\mathrm{OMS}}(t) = \ddot{n}_{\mathrm{OMS}}(t) + \ddot{n}_{\mathrm{GRS}}(t) > \ddot{n}_{\mathrm{OMS}}(t) \tag{4.9}$$

Hence, the PSD of the timeseries $(\Delta\ddot{x}_{12,\mathrm{OMS}} - \Delta\ddot{x}_{\mathrm{GRS}})$ is an upper limit to $S_{\Delta\ddot{x}_{12,\mathrm{OMS}}}$. The upper limit of the contribution to $S_{\Delta g_n}$ can then be evaluated as in Eq. (4.7). The result is shown in Fig. 4.9.

The information conveyed by the two methods is concordant, but measurements with grabbed TMs are much more significant and constraining.

We conclude that the low-frequency excess above Brownian is confidently ascribed to real forces, acting on the TMs, not associated with interferometer readout noise.

Between 40 µHz and 5 mHz, interferometer readout noise contributes at most to 1% of the total excess power.

4.3.2 Lever arm of excess noise

To have a better understanding of the nature of the low-frequency excess noise, we perform a cross-analysis with the angular acceleration time series, $\Delta\gamma_\phi$ and $\Delta\gamma_\eta$. The main reason why we perform this analysis is that electrostatic forces are applied with electrodes, off-centered with respect to the TMs. In particular, voltage noise on a single electrode would show up both as force and torque: a single electrode producing the entire noise would have an effective lever arm $|r_y| = 11\,\mathrm{mm}$, and $|r_z| = 0\,\mathrm{mm}$, generating torque along ϕ.
The idea is that the total measured force $f_x(t)$ could be caused by a distribution of forces with a distribution of application points, or possibly a single force with a single application point. In the general case, a distribution of forces would cause a total force and a total torque, defining an effective lever arm. In the single-force case, such a torque $\boldsymbol{N}$ would be associated with a lever arm $\boldsymbol{r}$: N_ϕ and N_η would be associated to

$$\begin{cases} N_\phi &= r_x f_y - r_y f_x \\ N_\eta &= r_z f_x - r_x f_z \end{cases} \tag{4.10}$$

meaning that the torque time series $N_\phi(t)$ and $N_\eta(t)$ would be correlated to the force time series $f_x(t)$, through the effective lever arms $(-r_y, r_z)$.

In Sec. 2.2.5, we noted that we have no access to the single-TM torque time series, but rather to the differential acceleration time series, which we defined in Eq. (2.10). Hence, we look for correlations among the measured Δg and $\Delta\gamma_\phi$, $\Delta\gamma_\eta$. We model the system as:

$$\begin{gathered} \Delta\gamma(t) = \Delta\gamma_0(t) + r\Delta\tilde{g}(t) \\ \Rightarrow\ \boldsymbol{\Sigma} = \begin{pmatrix} S_{\Delta\gamma_0}(f) + r^2 S_{\Delta\tilde{g}}(f) & r\,S_{\Delta\tilde{g}}(f) \\ r\,S_{\Delta\tilde{g}}(f) & S_{\Delta\tilde{g}}(f) \end{pmatrix} \end{gathered} \tag{4.11}$$

In this equation, we defined for simplicity $\Delta\tilde{g} = \Delta g\, M/I$. The quantity $\Delta\gamma_0$ represents the angular acceleration not correlating to Δg. The equation holds for both $\Delta\gamma_\phi$ and $\Delta\gamma_\eta$, with the proper definitions of r. In the second equation, we just give the frequency-domain CPSD matrix corresponding to the time-domain equation on the left. With this model, we can easily apply the results presented in Chapter 3 to infer r, which can be seen as the susceptibility of $\Delta\tilde{g}$ with respect to $\Delta\gamma$. We infer r employing the likelihood in Eq. (3.49), with the following assumptions:

1. We run the analysis on frequencies 36 µHz–462 µHz, using the first 7 frequencies, assuming that the lever arm r is frequency-independent.

2. We run the MCMC algorithm to infer the posterior distributions of $S_{\Delta\tilde{g}}(f)$ and r, assuming that r is real.

3. We assume that, at each frequency, the measured power spectrum $\boldsymbol{\Pi}(f)$ is complex-Wishart distributed, with theoretical CPSD $\boldsymbol{\Sigma}(f)$.

4. We perform this analysis on the set of 13 LTP noise-only runs.

Figure 4.10 shows the lever arms that we find with this analysis, r_z and r_y. The grey spans correspond to the average lever arms, weighted on the number of periodograms used for estimation. Respectively, the average effective lever arms are $r_y = (1.13 \pm 0.15)\,\mathrm{mm}$ and $r_z = (0.5 \pm 0.8)\,\mathrm{mm}$. The precision on r_y is much better, since noise in $\Delta\gamma_\phi$ is much lower than in $\Delta\gamma_\eta$. However, we note that r_y is far from the "electrical figure" 11 mm. From this, we can confidently state that the excess noise is not caused by single-electrode noise.

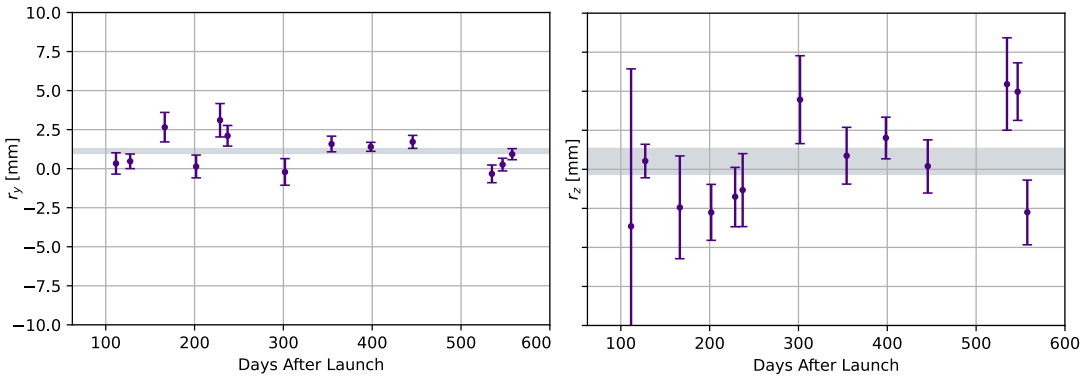


Figure 4.10: Effective lever arm r_y and r_z of the measured excess, inferred from the cross-correlations with the angular acceleration time series $\Delta\gamma_\phi$, $\Delta\gamma_\eta$. The grey spans, $r_y = (1.13 \pm 0.15)\,\mathrm{mm}$ and $r_z = (0.5 \pm 0.8)\,\mathrm{mm}$, represent the average lever arms, weighted with the number of periodograms used for estimation.

4.3.3 Mechanical stress

We conclude this chapter with a slightly off-track observation. As we mentioned in the introduction, any mass on the spacecraft is gravitationally coupled to the TMs: thus, any time-varying displacement of structural components is capable of inducing time-varying forces to the TMs. From this perspective, the effect would induce a force at the same frequency as the time-varying displacement. However, we noticed a quasi-DC variation of the instrument distortion, which we find important to report.

As said in the introduction, the quantity $\Delta x = (\Delta x_{12,\mathrm{GRS}} - \Delta x_{12,\mathrm{OMS}})$ is a measure of the GRS distortion. From the beginning of operations till the cold runs, i.e. before April 2017, quasi-static deformations in experimental runs were always detected. For example, in run Feb17B the distance between the two EHs expanded by 0.1 µm over 13.3 days. This is indeed an expected behavior, ultimately caused by the thermal expansion of the struts holding the GRSs. Indeed, the time series Δx shows a very high correlation with the average temperature. For example in Fig. 4.11, left panel, we show Δx and temperature during run Feb17B.

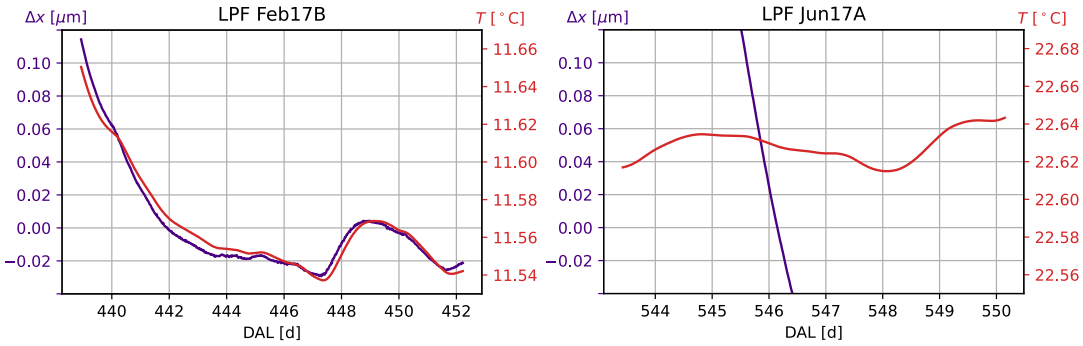


Figure 4.11: Quasi-static variation of the GRS distortion, measured as $\Delta x = (\Delta x_{12,\mathrm{GRS}} - \Delta x_{12,\mathrm{OMS}})$, and temperature, comparison between runs Feb17B and Jun17A. Scales are chosen to be the same, just offset, so that the comparison is immediate. Indeed, run Jun17A showed a strong, non-thermal, Δx drift.

After the cold runs, something different happened. Run May17E shows a 2.3 µm GRS expansion in 4.3 d, much faster than in February. Moreover, that expansion is not thermal, i.e. it is not correlated with temperature. In the following run Jun17A, the phenomenon is still evident (see Fig. 4.11, right panel), but, within Jun17B, it sets back to the nominal thermo-mechanical expansion.

It is likely that such a non-thermal expansion has been caused by the mechanical stress that involved the cold runs. Possibly, it could even be related to the mechanical event that triggered the LPF safe mode after run May17C. Indeed, in run May17D, the GRSs shrunk by 4.5 µm in just 2.8 d. However, in those runs it is not possible to estimate the correlation with temperature, as the EH thermometers were not available. Another similar non-thermal relaxation has been observed in run Sep16E.

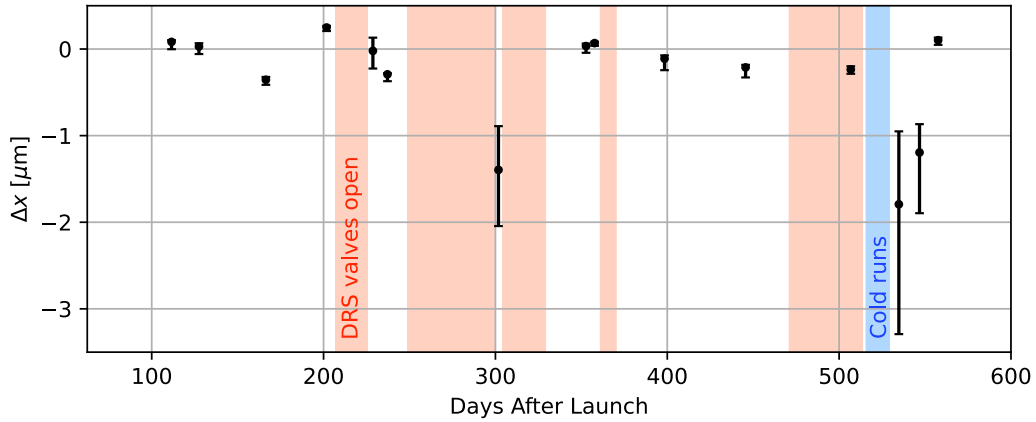


Figure 4.12: Mean quasi-static variation of the GRS distortion, measured as $\Delta x = (\Delta x_{12,\mathrm{GRS}} - \Delta x_{12,\mathrm{OMS}})$, in the 14 long-lasting runs. Points represent the mean variation, and spans show the minimum and maximum values reached during runs.

In Fig 4.12, we show the variation of the EH positions within runs. In Chapter 6, we will include the instrument distortion $\Delta x(t)$ in the analysis of in-band correlation among time series. We anticipate that the induced gravitational force is not a dominant effect on the overall measured acceleration noise. We also anticipate that we do not detect any relation between the in-band contribution to noise and the quasi-DC drift described here.

5 Brownian noise and outgassing environment

Brownian noise plateau, quasi-static acceleration drift, and pressure-related phenomena

Studying the evolution of the Brownian force noise is paramount in order to understand the evolution of pressure within the vacuum chambers. LPF did not have pressure gauges inside the VEs, because of their size and their impact on science measurements themselves. For the same reasons LISA will not carry pressure gauges, as they would not be essential to perform the science measurements but, rather, could influence the performance itself.

Nevertheless, LPF provided indirectly a measurement of pressure, via the differential acceleration measurements. Especially in the lowest-noise runs, the acceleration PSD shows a quite clear white-noise plateau at $\sim$ mHz frequencies, independent of frequency. The amplitude of such a PSD plateau evolves in time during the mission, decreasing in time. The reason why the plateau is so interesting is that its amplitude is strictly related to the pressure in the vacuum chambers [22]. The white-noise plateau, hence, effectively represents a proxy of pressure.

It has already been shown [2, 3] that the amplitude of such a plateau evolved in time during the LPF mission, in agreement with the expectation of pressure evolution. As we show in this chapter, we can confidently state that the measured white noise is largely contributed by the Brownian noise. Hence, we often refer to the white-noise plateau as the Brownian plateau. At $\sim$ mHz frequencies, we notice that the Brownian plateau is the limiting source of noise.

In this chapter, we analyze the evolution of the Brownian PSD level, providing an estimate of the pressure within the vacuum chambers, and a projection for LISA. We analyze the pressure evolution, analyzing and determining its dependence on time and temperature. The evolution of pressure yields details about the outgassing environment within the vacuum chambers, and allows to better understand which sources contribute the most to the overall pressure. We also analyze to which extent the noise plateau is ascribable to Brownian noise, especially during low-temperature runs and DRS runs. We also analyze the relations with the long-term Δg drift.

5.1 Brownian motion and pressure

Brownian noise is the manifestation of the random forces given to the TMs by the random impacts with the gas molecules within the experimental apparatus. It is intrinsically an inevitable effect, ultimately due to the exchange of momentum between the residual gas molecules and the test masses. Such noise affects all the degrees of freedom: the translational x, y, z, and the angular ϕ, η, θ. The effect is quite faint in absolute terms, so that, even in the high-sensitivity LPF case, it is visible only in interferometric measurements. In our case, as will be clearer later, it is well-detectable only in the sensitive-axis x direction. Along the other interferometric coordinates, ϕ and η, it is quite indiscernible.

In [22, 28], it is shown that Brownian motion appears as white, frequency-independent noise, both in the force and torque PSD. In particular, [22] shows that, for a single LISA TM, the force (torque) PSD plateau is given by

$$S_F = 4k_B T \beta_{\mathrm{tr}} \qquad S_N = 4k_B T \beta_{\mathrm{rot}}, \tag{5.1}$$

where β_{tr} and β_{rot} are respectively the translational and rotational damping coefficients, mainly depending on the system geometry, the gas species, and pressure. These coefficients, in the same references, have been calculated and experimentally measured.

Some points might be relevant for further discussions:

1. The damping coefficients can be calculated, under the assumption that the TMs are perfect cubes surrounded by an infinite volume of gas. Even though this approximation is not valid in the LISA/LPF configuration, these calculations show that both coefficients are *a)* proportional to pressure P, *b)* proportional to inverse square root of temperature $T^{-1/2}$, *c)* proportional to the square root molecule mass $m^{1/2}$.

2. In [22], the damping coefficients have been experimentally measured, with torsion pendulums at the University of Trento, using LISA-like geometries closely resembling the flight models. The same reference also validated the results with simulations. The rotational coefficient has been measured using a torsion pendulum with 1 TM; analogously, a combination of the translational and rotational coefficients has been measured with a cross-shaped 4-TM pendulum. Measurements have been performed at various pressure values, to isolate the coefficients from their dependence on pressure. Nevertheless, measurements have been performed at a stable temperature $(293.0 \pm 0.1)\,\mathrm{K}$, with no variations. Given the LPF temperature variation and the assumed theoretical dependence on temperature, the coefficients are accurate to within 5% precision. However, the coefficients are much more sensitive to the gas species' molar mass: being proportional to $m^{1/2}$, at the same pressure and temperature, the damping coefficient for air is 25% higher than that of water gas, and that for hydrogen is 75% lower.

3. We separate the contribution of the linear and angular damping coefficients in measurements in [22], to estimate their value in the LPF flight conditions.

Following the measurements in Table I in [22], we write the following equations

$$\begin{cases} \beta'_{\mathrm{IV}} &= (r^2\beta'_{\mathrm{tr}} + \beta'_{\mathrm{rot}}) + 3(r^2\beta'^{\infty}_{\mathrm{tr}} + \beta'^{\infty}_{\mathrm{rot}}) - 4(r^2\beta'^{\infty}_{\mathrm{tr}} + \beta'^{\infty}_{\mathrm{rot}}) \\ \beta'^{\infty}_{\mathrm{III}} &= 4(r^2\beta'^{\infty}_{\mathrm{tr}} + \beta'^{\infty}_{\mathrm{rot}}) \\ \beta'^{\infty}_{\mathrm{I}} &= \beta'^{\infty}_{\mathrm{rot}} \\ \beta'_{\mathrm{I}} &= \beta'_{\mathrm{rot}} \end{cases} \tag{5.2}$$

where the subscripts I, II, III, IV, refer to the columns in Table I in [22]. The superscript ∞ refers to the theoretical value, with the TM not surrounded by the close-by GRS. Values with no superscript indicate the case with the TM enclosed in the GRS. The superscript $'$ indicates the partial derivative with respect to pressure, for clarity. In the first equation, we have also considered that in the 4TM torsion pendulum only one TM is surrounded by the GRS.

These equations lead to the following values:

$$\partial\beta_{\mathrm{tr}}/\partial P = 5.32 \times 10^{-4}\,\mathrm{m\,s} \qquad \partial\beta_{\mathrm{rot}}/\partial P = 4.8 \times 10^{-8}\,\mathrm{m^3\,s} \tag{5.3}$$

We recall that these values hold at a temperature 293 K, with the LPF EH geometry, assuming that the outgassing species is water gas. In the following, we only focus on the translational damping coefficient, which is of interest for the LPF noise performance.

4. The damping coefficients translate into force and torque noises. In the LPF configuration, noises on the two TMs are completely independent, as the TMs outgassing environments are isolated. We focus on the contribution to force. In the differential measurements, the two TMs contribute equally to the total PSD, following

$$S_F = \frac{1}{2}M^2 S_{\Delta g}, \tag{5.4}$$

where M is the TM mass. With the computed β_{tr}, the two-TM differential PSD would be $(2.1 \times 10^{-15}\,\mathrm{m\,s^{-2}\,Hz^{-1/2}})^2/\mu\mathrm{Pa}$. This value, as noted in the reference, is compatible with the value obtained from simulations, $(1.8 \times 10^{-15}\,\mathrm{m\,s^{-2}\,Hz^{-1/2}})^2/\mu\mathrm{Pa}$.

5. Following [22], we use the value obtained from simulations, for the conversion factor between the Brownian acceleration PSD and pressure. Ultimately, this is the value that we use:

$$\partial S_{\Delta g}/\partial P_{\mathrm{H_2O}} = \left(1.8 \times 10^{-15}\,\mathrm{m\,s^{-2}\,Hz^{-1/2}}\right)^2/\mu\mathrm{Pa} \tag{5.5}$$

with $P_{\mathrm{H_2O}}$, we explicitly indicate that the coefficient assumes that the emitted gas is water.

5.2 LPF linear Brownian noise

The white plateau PSD, as measured in LPF's long ordinary runs (orange in Table A.2), is shown in Fig. 5.1, as purple points. Attached to every point, a label indicates the mean temperature of the run. Values and errors have been calculated with the MCMC algorithm described in Sec. 4.1. With that calculation method, we recall that the Brownian posterior is essentially that of the lowest measured PSD, with an associated error that is basically driven by the number of averaging windows used for its computation[1]. We also recall that the Brownian estimate does not depend substantially on the method used for its evaluation, meaning that the values in the figure are pretty robust[2]. The reason is that, at $\sim$ mHz frequencies, the number of averaging windows for PSD computation is high, of the order of 10^3, making the estimation solid.

In what follows, we fit the evolution of the Brownian plateau with expected models. Before doing that, we note a few relevant points:

1. The Brownian level (purple points) decreases in time across the whole LPF mission, apparently linearly in a log-log plot, at least in the runs at nominal 22 °C temperature. No apparent apparent saturation values are evident. The Brownian level, however, does show some fluctuations, in particular it is higher in runs with a higher mean temperature (in July 2016, the mean temperature was 24 °C).

2. Lowering temperature to 11 °C reduced the Brownian level, as expected since it reduces the outgassing rate.

3. Lowering temperature to 0 °C also reduced the white-noise plateau, down to $S_B = (1.7 \pm 0.1)\,\mathrm{fm^2\,s^{-4}\,Hz^{-1}}$, which is the lowest white-noise plateau measured on LPF. This value has been obtained using glitch-free stretches during the longest cold run May17B [3]. However, as we will notice later, this value is not as low as it would be expected if it were entirely due to Brownian noise. The lowest Brownian plateau measured on LPF, which is ascribable to Brownian noise, has been measured in run May17E, $S_B = (2.54 \pm 0.15)\,\mathrm{fm^2\,s^{-4}\,Hz^{-1}}$. According to the conversion factor in Eq. (5.5), it corresponds to a (water) pressure of $P_{\mathrm{H_2O}} = (0.78 \pm 0.05)\,\mu\mathrm{Pa}$.

4. The Brownian measured during the DRS run Apr17A, about $4\,\mathrm{fm^2\,s^{-4}\,Hz^{-1}}$, is higher than the previous run Feb17B, despite being measured two months later. This shows that some other mechanisms influenced the white level in run Apr17A.

[1] The number of periodograms available at mHz frequency is of the order of 10^3, making the inverse-gamma PSD distribution (Eq. (3.18)) basically Gaussian, with mean Π and variance Π/M.

[2] Brownian levels in this Section, as shown in Fig 5.1, slightly differ from the ones in [5]. Consequently, parameters Θ and γ are slightly different, but still compatible within 1σ. The reason is that in [5] we use spectral windows a few samples shorter, leading to slightly different PSD values. Conclusions and physical interpretations are, however, unchanged.

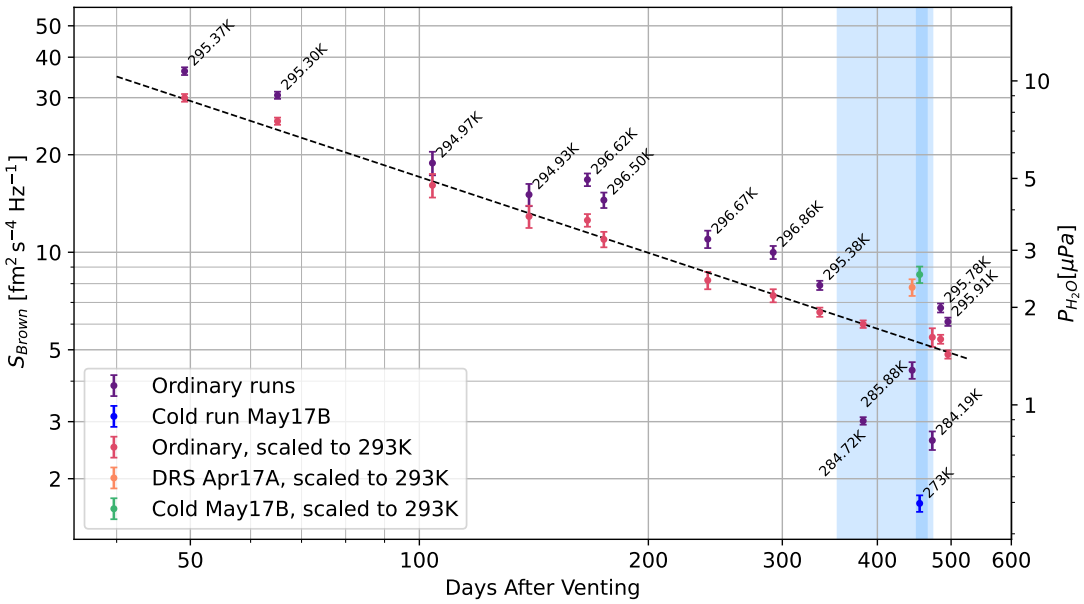


Figure 5.1: Brownian plateau evolution during the LPF mission, versus the days after venting (62.4 DAL). On the left axis, the white noise PSD plateau; on the right axis, the corresponding pressure value assuming water outgassing, according to Eq. (5.5). The purple points represent Brownian during the long ordinary runs of Table A.2, and the labels indicate the mean temperature of the runs. The light blue region indicates runs performed at temperatures lower than nominal, including runs Feb17B and May17E at 11 °C. The darker blue region, and the blue point within it, represent the measured Brownian plateau during run May17B, at nominal 0 °C. The red, orange, and green points represent the Brownian values rescaled to 293 K, respectively of ordinary runs, the DRS run Apr17A and the cold LTP run May17B. Rescaling is performed as described in text, and corresponds to an activation temperature $\Theta = (6.9 \pm 0.2)$ kK. The dashed line is a fit, based on the outgassing model described in text.

5.2.1 Outgassing rate and Brownian rescaling to 293 K

As computed in [28], and as validated with experiments and simulations, the damping coefficient is proportional to pressure. Hence, the Brownian PSD is proportional to pressure. If the outgassing element is water, then the coefficient has been estimated in Eq. (5.5); in any case, the proportionality between Brownian level and pressure holds. Here, we use the Brownian level as a proxy for pressure.

Let's now work out some well known results of vacuum systems. The commonly used vacuum pumping equation [67–69] states that, in a small-volume vacuum system pumped at fixed temperature T, the following simple equation is satisfied:

$$P(t) = \frac{Q(t)}{S}, \tag{5.6}$$

where $P(t)$ is the system pressure; S is the pumping speed, depending on the used conductances and vacuum pumps; $Q(t)$ is the outgassing rate. The outgassing rate is not a specific property of materials, even though materials can usually be associated to an order-of-magnitude estimate of their outgassing rate, based generally on their capability of adsorbing particular gas species. Thermal treatments and coatings, as well as the surface history, may strongly affect the outgassing rate.

As said in Chapter 2, the characteristic GRS's pumping speed from the inner EHs to the outer space was, for water gas, about 19 L/s. Equation (5.6), however, needs some adaptation to the GRS case. The GRS is not a small-volume system, hence the equation would generally change into the following convolution:

$$P(t) = \frac{(Q * P_\delta)(t)}{S} = \frac{1}{S}\int_0^t Q(\tau) P_\delta(t-\tau)\, d\tau, \tag{5.7}$$

where $P_\delta(t)$ is the system response function to an outgassing impulse, whose shape depends on the system geometry. Nevertheless, the GRS volume order of magnitude is 10 L: this, with the pumping speed above, yields a typical time scale of about $\tau_d \sim V/S \sim 0.3$ s, which is the time scale of $P_\delta(t)$. This time scale is much shorter than the variation of Brownian over the mission time scale (tens of days). Therefore, the latter equation can effectively be approximated with the previous Eq. (5.6), from which we deduce that pressure is proportional to the outgassing rate. Hence, the dependence on Brownian on time and temperature has to be searched in the dependence of the outgassing rate on time and temperature.

We add another well-known property of vacuum systems: the desorption of gaseous elements from walls commonly follows an Arrhenius equation, associated with an activation temperature Θ [67, 68, 70]. Therefore, a higher temperature enhances the process, yielding higher outgassing. The activation temperature is an intrinsic property of both the wall material and the outgassing species, and in our case it will be used to understand which element is being emitted.

The dependence of outgassing on temperature and time can not be separated, in principle. It is indeed straightforward to think that an aggressive baking process with high temperature would extract gas from the outer wall surface, leaving fewer adsorbed particles than before the process. After restoring the previous temperature, the system

would not behave as if temperature never changed. In our case though, during the LPF mission, the GRS was not subject to great temperature changes. Hence, we assume that changing and restoring temperature brings the system back to its original outgassing evolution function. Mathematically, this means that temperature and time are separable variables:

$$Q(T,t) = e^{-\Theta/T} Q(t) = e^{-E_a/RT} Q(t) \tag{5.8}$$

In this equation, Θ is the activation temperature, E_a the activation energy, and R the gas constant $R \sim 8.31\,\mathrm{J\,mol^{-1}\,K^{-1}}$. The two expressions are equivalent, as the activation energy is $E_a = RT$; the activation energy is usually expressed in MJ/kmol.

Applying Eq. (5.8) to the LPF Brownian evolution, we rescale the measured Brownian levels to the same temperature, which we choose 293 K. This process allows us to separate the dependence of Brownian on temperature and time, and to proceed as if the system was always kept at constant temperature. We rescale Brownian as:

$$S_{B,293\,\mathrm{K}}(t) = e^{\left(\frac{\Theta}{T} - \frac{\Theta}{293\,\mathrm{K}}\right)} S_B(t) \tag{5.9}$$

The activation temperature Θ is not known a priori. We will infer it in the next paragraph after modelizing the time evolution $S_B(t)$. However, we anticipate that we find an activation temperature of 6.9 kK. Using this value in Eq. (5.9), and applying it to the measured Brownian levels in ordinary runs, we find the red points in Fig. 5.1. Applying it to the measured Brownian levels in the DRS run Apr17A and the cold run May17B, we find the values represented in orange and green in Fig. 5.1.

5.2.2 Outgassing rate evolution

After finding a model of the temperature dependence of the outgassing rate, we need to find a model for its time evolution function. A commonly-used model, very often measured experimentally, is that of the power-law dependence of the outgassing rate on the pumping time. Let's denote with t_v the starting time of the pump-down, which in our case is the time of venting to outer space. The model assumes that the outgassing rate depends on time t, as $(t - t_v)^{-\gamma}$. The exponent γ is a property of the emitting material and gas species: generally, it may vary between 0.5 and 1, but also higher exponents, 3-4, might be possible [68, 69], see later discussion.

Therefore, we model the Brownian evolution as

$$S_B(T,t) = A\, e^{\left(\frac{\Theta}{T} - \frac{\Theta}{293\,\mathrm{K}}\right)} \left(\frac{t_v}{t - t_v}\right)^{\gamma} \tag{5.10}$$

Here, $t_v = 62.4\,\mathrm{d}$ is a fixed value, chosen as the venting to space of the two GRSs. The exponent γ, the activation temperature Θ, and A, are fitting parameters. The varying quantities, t and T, are chosen as the mean time and temperature of the runs taken into account. We notice the presence of an overparametrization in the numerator, $A\, t_v^{\gamma}$: it is done so that A and S_B have the same units, and it does not change the result, as A has no particular physical meaning. The relevant parameters, in this model, are Θ and γ, which might hint at the outgassing mechanisms occurring.

We perform the fit with a least-squares method, implementing it with an MCMC algorithm to find possible correlations between parameters, and to be sure that the system does not present any bimodalities. For reasons that will be clear later, we just perform the fit over the ordinary runs, that is, for the red points in Fig. 5.1. The best-fit result is depicted in the same figure as a dashed line. Parameter histograms and correlations are shown in Fig. 5.2.

In particular, we find the activation temperature $\Theta = (6.9 \pm 0.2)\,\mathrm{kK}$, and the decay power-law exponent $\gamma = (0.78 \pm 0.02)$. The activation temperature corresponds to an activation energy $(57 \pm 2)\,\mathrm{MJ/kmol}$. Parameters A and γ show a slight positive correlation, which, as noticed before, can be ascribed to the overparametrization of the numerator.

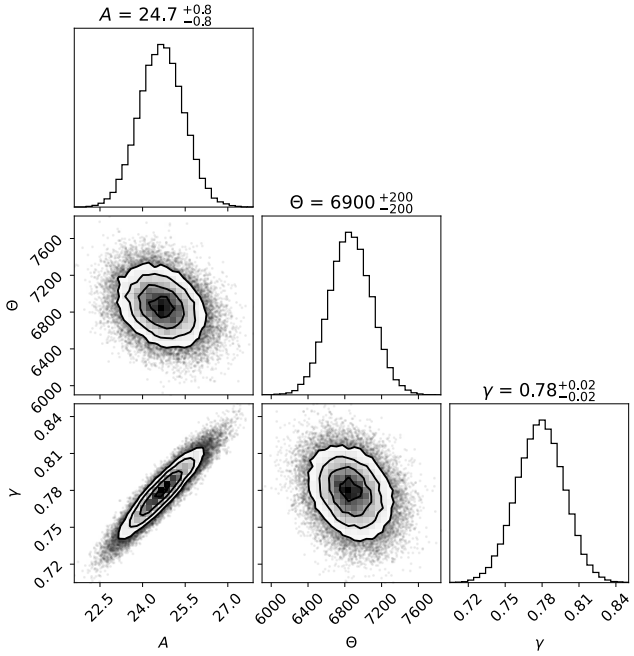


Figure 5.2: Parameters from the fit of the measured Brownian levels in ordinary runs (purple points in Fig. 5.1, except for the DRS run), with the model in Eq. (5.10). A line with the best fit is represented as a dashed line in Fig. 5.1, represented as a fit to the temperature-scaled Brownian data. In the histograms, a slight positive correlation between A and γ might be noticed, which is ascribable to the model definition and has no physical importance or meaning.

Fit results

We note a few points about the fit results:

Re-scaling. The measured Brownian data, during ordinary runs, are well fitted to the model in Eq. (5.10), as can be seen in Fig. 5.1. Such ordinary runs include remarkably two runs performed at 284 K, Feb17B and May17E, whose Brownian levels follow the model, despite being performed at a temperature 11 K lower than nominal. The same fact can be noticed in some runs performed at 296 K, 2 K above nominal, for which the temperature re-scaling effectively brings the Brownian levels to the best-fit power-law line.

DRS run. Run Apr17A, performed at 285 K as run Feb17B, but performed with DRS thrusters, shows a Brownian level higher than the previous run, despite being performed two months later. If re-scaled to 293 K, the Brownian level is about twice the expected value (orange point). This means that another mechanism, possibly related to the DRS thrusting mechanism, is inducing a non-Brownian white noise. If we estimate the Brownian-related white noise following the best-fit line, the excess white noise over Brownian can be estimated at $2.6\,\mathrm{fm^2\,s^{-4}\,Hz^{-1}}$.

Cold run. The same occurs for the cold run performed at nominal 273 K. On the one hand, it shows a lower Brownian level with respect to ordinary runs performed at a higher temperature (dark blue point in Fig. 5.1), showing that Brownian during ordinary runs did not saturate. On the other hand, performing the temperature re-scale shows that the measured value is however higher than expected. The excess white noise can be estimated at $0.7\,\mathrm{fm^2\,s^{-4}\,Hz^{-1}}$.
The origin of this excess white noise is not certain. Still, it is likely that, despite glitch-free stretches being chosen for Brownian evaluation [3], some fast low-impulse events are still included in the stretches, increasing the white-noise level.

Excess white noise. Ordinary runs performed at 284 K, especially run Feb17B which has a high precision due to a long measurement time, can be used to estimate the presence of an additional non-Brownian noise level, given that their Brownian is particularly low. We note that, after temperature re-scaling, Brownian in run Feb17B is on the best-fit line. If additional white noise was present (not scaling with temperature), common to all runs, it could be at maximum $0.5\,\mathrm{fm^2\,s^{-4}\,Hz^{-1}}$, not to shift the observed values further than 3σ from the best-fit line. Therefore, the observed excess in runs Apr17A and May17E are not ascribable to an excess common to all runs.

Venting time. We chose the venting time t_v as a fixed parameter in the fit to the model. However, it could not be the case. It is true that the vacuum chamber valves were opened 62.4 DAL, but, if the outgassing sources were located between the valves and the venting ducts, then the venting time should be chosen as the launch time 0 DAL. We performed an additional fit, leaving t_v as a free parameter. The fit is compatible with data, and the result is that $t_v = (63 \pm 8)$ DAL, compatible with the venting time 62.4 DAL. Parameters Θ and γ are compatible with the previous ones. This confirms that the major source of outgassing is located inside the vacuum chambers.

5.3 Long-term drift and pressure unbalance

A short digression is necessary to better characterize the outgassing environment and gas pressure on LPF's TMs. We show that the long-term drift detected in $\Delta g(t)$ is possibly related to a small pressure unbalance between pressure on the two TM sides. The mechanism that relates the two quantities is rather articulated, and requires a little introduction.

It is known [1–3] that, in addition to the measured quasi-stationary noise, LPF Δg time series were affected by a long-term drift. This drift showed up as a steady variation of the differential force Δg, with a magnitude of the order of $\partial \Delta g_{\rm DC}/\partial t \sim 1\,\mathrm{pm\,s^{-2}\,d^{-1}} \sim 10\,\mathrm{am\,s^{-3}}$. This drift, being quasi-DC, did not influence the PSD described in the previous sections. A first reason is that it only affects frequencies lower than the analyzed ones; moreover, our PSD estimation method allows to reduce the influence of the DC-frequency bin on PSD estimates.

Contributions to long-term drift

Some mechanisms are known to contribute directly to long-term drift. First, we list them, and subtract the known ones from the $\Delta g(t)$ time series, so that only the residual drift is left.

1. **Propellant gravitational field.** The major source of DC drift is the variation of the gravitational attraction by propellant tanks, due to the depletion of cold nitrogen. As described in Sec. 2.1.4, the contribution from this source is known with sufficient precision, as the gas mass is known from the flowmeter measurements and the tank gravitational field is calculable. However, we note that, despite the relatively high precision, measurements could be affected by systematic errors due to calibration uncertainty, as thrusters are used in-loop with TM position measurements. We assume that the fuel contribution $\Delta g_{F,\rm DC} = \kappa_t\,\kappa_b\,\Delta g_{F,\rm DC,t,b}$, depending on two calibration factors, κ_t and κ_b, respectively representing calibration corrections for the used tanks and fuel branch. We find the calibration correction *a posteriori*, as explained later, with the fit provided later in Fig. 5.6. We find that κ_t is: $[\kappa_1 = 1.00,\ \kappa_2 = 1.05,\ \kappa_3 = 0.95]$; we find that κ_b is $[\kappa_A = 1.00,\ \kappa_B = 0.92]$. These values agree with expectations: calibration during station-keeping maneuvers was estimated to be precise with 10% uncertainty.

2. **Thermal effects.** Thermal effects contribute directly to acceleration PSD. As shown in later chapters, temperature and Δg are highly correlated at low frequencies, where temperature fluctuates the most. At such frequencies, electronic noise in temperature readouts is negligible. Moreover, we note that including temperature in our fit for DC drift also allows to infer the thermal "susceptibility" coefficient.

3. **GRS distortion.** Instrument distortion, as described in Sec. 2.2.6, contributes to Δg, coupling it to the time series $(\Delta x_{\rm GRS} - \Delta x_{\rm OMS})$ through a known stiffness ω_d^2. The major mechanism causing instrument distortion is thermal, hence it

is highly coupled to the temperature time series. However, the time series $-\omega_d^2(\Delta x_{\rm GRS} - \Delta x_{\rm OMS})$ is precisely known, and can be subtracted first. We note that the contribution is, however, rather small.

4. **TM charging.** Test mass charging in time, linearly varying because of the interaction with cosmic rays, causes a steady drift in the resulting force. The induced differential acceleration drift reads

$$\Delta\dot{g}_{\rm DC} = \frac{\partial q_2/\partial t}{M\,C_T}\left|\frac{\partial C_x}{\partial x}\right|\Delta_{x,2} - \frac{\partial q_1/\partial t}{M\,C_T}\left|\frac{\partial C_x}{\partial x}\right|\Delta_{x,1}. \tag{5.11}$$

Only the second term is relevant since, as shown in Table A.5, $\Delta_{x,2}$ has always been compatible with 0 during the mission. On the contrary, $\Delta_{x,1}$ required active compensation. It was not compensated at the beginning of the mission, amounting to $\Delta_{x,1} \sim -24\,\mathrm{mV}$. Estimating Eq. (5.11) at the beginning of the mission, we find $\Delta\dot{g}_{\rm DC,ch} = 0.4\,\mathrm{am\,s^{-3}}$, which is negligible with respect to the residual DC Δg drift.

In Fig. 5.3, we show as example the contribution from propellant gravitational attraction, thermal effects, and GRS distortion to the DC drift in run Sep16E, lasting 2.8 days. Propellant gravitational attraction contribute to 78% of the total DC Δg drift.

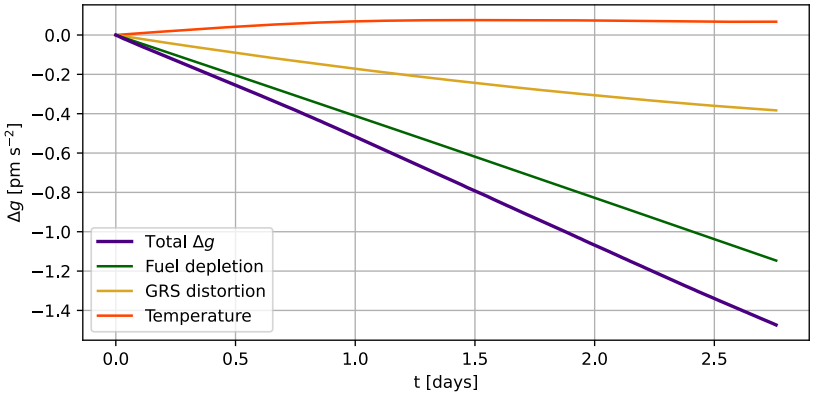


Figure 5.3: Long-term evolution of Δg during run Sep16E, lowpassed, and shifted so that all time series start at the axes origin. Drift due to fuel depletion accounts from 78% of the total drift slope.

We fit data as follows. First, for each run, we subtract from the Δg time series the known time series, i.e. the contribution from fuel depletion Δg_F and the contribution from GRS distortion $\Delta g_{\rm dist}$. We define $\Delta g_c(t) = \Delta g(t) - \kappa_t\kappa_b\Delta g_F(t) - \omega_d^2\Delta g_{\rm dist}$. Then, for each run, we lowpass data with a 100-s Blackman-Harris sliding window, to remove the high-frequency noise; then we partition data into 1-day long, non-overlapping segments. Finally, we fit data from each stretch to the following model, which allows to find the temperature susceptibility α_T and the average DC differential drift $\Delta\dot{g}_{\rm DC}$. The last term, c_0, is a constant that has no influence on our analysis.

$$\Delta g_c(t) = \alpha_T T(t) + \Delta\dot{g}_{\rm DC}\,t + c_0 \tag{5.12}$$

We perform the fit with a least-squares method, from which we extract the best-fit parameters. For each run, we average parameters and compute their variances, the latter computed as the standard deviation of the best-fit parameters among the 1-day segments. As some runs contain a low number of averaging segments, we level out the variances, so that the variance of parameters in the k-th run, computed with M_k segments, is

$$\sigma^2_{\alpha_T,k} = \frac{\sum_{k=1}^{13}(N_k - 1)\sigma^2_{0,\alpha_T,k}}{\sum_{k=1}^{13}(N_k - 1)} \frac{1}{N_k} \tag{5.13}$$

where $\sigma^2_{0,\alpha_T,k}$ is the variance evaluated previously. The evaluation of $\sigma^2_{\Delta\dot{g}_{\mathrm{DC}},k}$ follows the same scheme.

Residual long-term drift evolution

In Figure 5.4, we show the results of the fit described in the previous paragraph. In the left panel, we show the evolution during ordinary runs of the acceleration drift. On the right panel, we show the evolution of the thermal coefficient α_t. Regarding the latter, we underline that it is consistent with the values found in [71], and with the values from time series decorrelation, inferred in Chapter 6.

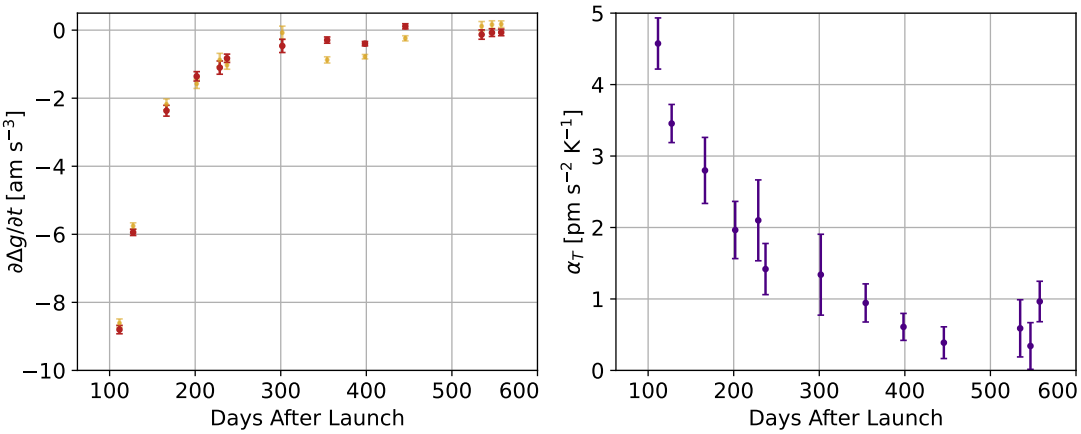


Figure 5.4: Residual differential DC drift slope (left panel), and temperature susceptibility coefficient (right panel), as resulting from the fit to data, according to the model in Eq. (5.12). Red points in the left panel represent data with fuel mass loss calibration corrections κ_t, κ_b. Dark yellow points represent non-corrected data $\kappa_t = \kappa_b = 1$.

We note that both $\Delta\dot{g}_{\mathrm{DC}}$ and α_T decay to zero in their evolution. This behavior is expected for α_T, since thermal effects decay with pressure, which decays in time. However, in principle, it could be unexpected for $\Delta\dot{g}_{\mathrm{DC}}$. This hints that the two parameters might have some sort of relation, between them and with pressure.

We find, indeed, a rather good correlation between α_T and S_B, but we do not find a significant correlation between $\Delta\dot{g}_{\mathrm{DC}}$ and S_B. Nevertheless, we find two additional correlations, which might hint at a pressure-related mechanism, as described in the following paragraph.

Correlations and pressure unbalance

First, we find a significant linear relation between α_T and the derivative $\partial S_B/\partial T$. Following from Eq. (5.8) and the proportionality between Brownian level and outgassing, $\partial S_B/\partial T = (\Theta/T^2)S_B$. In Figure 5.5, we plot the two quantities and show the linear fit. The linear fit slope is $\xi_T = (1.48 \pm 0.07) \times 10^{18}\,\mathrm{s\,m^{-1}}$, with p-value ~ 0.6.

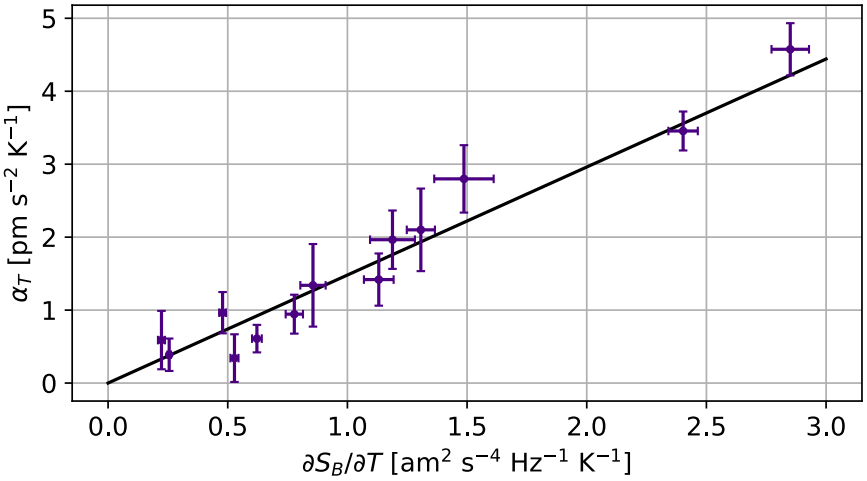


Figure 5.5: Plot showing a linear relation between $\alpha_T = \partial\Delta g/\partial T$ and the derivative $\partial S_B/\partial T$. The linear fit slope is $\xi_T = (1.48 \pm 0.07) \times 10^{18}\,\mathrm{s\,m^{-1}}$.

The second correlation that we find is between $\Delta\dot{g}$ and the time derivative $\partial S_B/\partial t$. Following from Eq. (5.10), $\partial S_B/\partial t = -(\gamma/(t - t_v))S_B$. In Figure 5.5, we plot the two quantities and show the linear fit. The linear fit slope is $\xi_t = (1.67 \pm 0.04) \times 10^{18}\,\mathrm{s\,m^{-1}}$. The p-value, however, is not optimal. We find it realistic and very likely that the quality of the fit is not excellent because of calibration errors of the flowmeters. As noted before, drift due to fuel depletion is dominant in the Δg time series, therefore even a small calibration error in fuel mass loss can lead to large variations in the residual drift. We choose to use the fit in Fig. 5.6 to find, *a posteriori*, optimal values for κ_t and κ_b, optimizing the fit goodness. The resulting κs are compatible with 1 within 10%, consistently with calibration results. Therefore, we find this procedure adequate. Even if the κs were slightly different, conclusions and physical interpretations would still remain unchanged.

For completeness, Fig. 5.4 on the left panel, and 5.6, also show (light points) the results with $\kappa_t = \kappa_b = 1$. The fit slope in Fig. 5.6 would be $(1.70 \pm 0.04) \times 10^{18}\,\mathrm{s\,m^{-1}}$, and the difference with the line in Fig. 5.6 would not be visible by eye.

We highlight that, within uncertainties, $\xi_T = \xi_t$, which we call ξ. We point out that ξ_T and ξ_t have been found independently. This observation supports a proportionality between the DC value Δg_{DC} and the Brownian which, ultimately, hints at proportionality between Δg_{DC} and pressure. Recalling the proportionality between Brownian and (water) pressure in Eq. (5.5),

$$\Delta g_{DC} = \xi\, S_B + \Delta g_{0,k} = \xi\, \frac{\partial S_{B,\Delta g}}{\partial P_{\mathrm{H_2O}}}\, P_{\mathrm{H_2O}} + \Delta g_{0,k}, \tag{5.14}$$

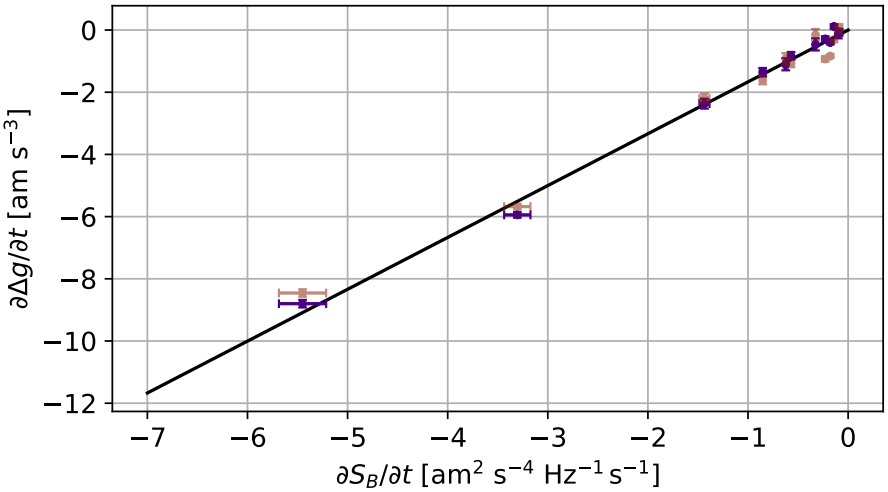


Figure 5.6: Plot showing a linear relation between $\Delta\dot{g} = \partial\Delta g/\partial t$ and the derivative $\partial S_B/\partial t$. The linear fit slope is $\xi_t = (1.67 \pm 0.04) \times 10^{18}\,\mathrm{s\,m^{-1}}$. Purple points represent data with fuel mass loss calibration corrections κ_t, κ_b. Dark yellow points represent non-corrected data $\kappa_t = \kappa_b = 1$.

where $\Delta g_{0,k}$ is an integration constant, which might be different in different runs, and does not affect our analysis.

A static acceleration Δg_{DC}, proportional to pressure $P_{\mathrm{H_2O}}$, can easily be interpreted as a static pressure unbalance, i.e. a constant ratio between pressure on the positive TM faces and the negative TM faces. This mechanism would give a reason why the static acceleration decreases along with pressure. In formulas,

$$\frac{\Delta P_{\mathrm{H_2O}}}{P_{\mathrm{H_2O}}} = \frac{\xi M}{L^2}\frac{\partial S_{B,\Delta g}}{\partial P_{\mathrm{H_2O}}} = (4.6 \pm 0.3) \times 10^{-3} \tag{5.15}$$

where M and L^2 are the TM's weight and face area. We note that, as LPF provided differential measurements, $\Delta P_{\mathrm{H_2O}}/P_{\mathrm{H_2O}}$ is to be intended as an average between the two GRS. Since the GRSs were geometrically nominally identical, however, it is straightforward to think that the difference could be small. We also notice that this ratio assumes that the outgassing element is water; if that is not the case, it would be slightly different.

5.4 Conclusions on the GRS outgassing environment

Summarizing the analyses presented in this chapter, we found a number of mechanisms relating residual gas pressure around the TMs to other measurable dynamical quantities. These mechanisms allow to analyze pressure evolution and the outgassing environment inside LPF's vacuum chambers, which would otherwise not be accessible, because of the absence of traditional pressure gauges.

Through the analysis of the Brownian white-noise plateau, we find it to be an excellent proxy of gas pressure and outgassing rate, for all ordinary runs. Small deviations from a power-law behavior before Feb17 are confidently ascribed to temperature variations; such behavior is confirmed also during the cooldown to 11 °C, and is not influenced by the cooldown to 0 °C. Cold runs and DRS runs, however, show an excess white noise, not ascribable to pressure effects. Rather, we find it likely that *a)* white noise in cold runs could be affected by the presence of unsubtracted glitches, and that *b)* white noise in DRS runs could be spoilt by the activation of the DRS thrusters.
By fitting the measured evolution of Brownian noise, we find some properties related to the outgassing environment.

1. We find that pressure follows the model

$$P(T,t) = Ae^{-\Theta/T}\left(\frac{t_v}{t-t_v}\right)^{\gamma}, \tag{5.16}$$

   with an activation temperature Θ and an exponent γ. Such a time dependence is often observed in vacuum systems, and could be modeled as desorption from a reversibly adsorbed surface monolayer.

2. The activation temperature $\Theta = (6.9 \pm 0.2)\,\mathrm{kK}$, corresponding to an activation energy $(57 \pm 2)\,\mathrm{MJ/kmol}$, is generally compatible with the desorption of water from metal.

3. The decay exponent $\gamma = (0.78 \pm 0.02)$ is also compatible with emission of water from metal. Usually, experiments yield γ values ranging from 0.7 to 2.0 [69], strongly depending on the emission mechanism. A single monolayer could be modeled to follow a power law $\propto 1/t$; however, a different exponent could be due to more complex desorption [68, 69], and it is not rare. If outgassing was due to diffusion from bulk, rather than a surface adsorbed layer, pressure would follow the law $\propto 1/t^{0.5}$. Therefore, an exponent of the order of 0.8 is understandable.

4. We find that the venting time, occurring on 62.4 DAL, is compatible with the starting time of the decaying power law. This hints that the major outgassing sources were located within the vacuum enclosures. Another option would be that sources were located outside the VE valve, but such sources would show a decay starting time occurring on DAL 0.

   A possible location for the outgassing sources could be the cables located inside each VE, which expose a large area in front of the x TM faces. Even though, in most cases, outgassing from polymers follow more complex evolution

functions, also showing power-law exponents higher than 2, measurements are still compatible with our hypothesis. Indeed, measurements on the LPF flight-model cables, performed at CERN, showed that an emission following a $\propto 1/t$ behavior is possible. Other options include the W IBM, the Mo EH, or a combination of them.

5. Finally, we find a correlation between the observed Brownian evolution and the observed long-term Δg drift, and provide a possible mechanism at the basis of such correlation. Indeed, the presence of a static pressure unbalance between the two x sides of the TMs would generate a static force decaying along with Brownian noise, as Brownian is proportional to pressure.

   To support this hypothesis, we find two relations showing direct proportionality: $\partial \Delta g_{\mathrm{DC}}/\partial t \propto \partial S_B/\partial t$ and $\partial \Delta g_{\mathrm{DC}}/\partial T \propto \partial S_B/\partial T$, which point at $\Delta g_{\mathrm{DC}} \propto \partial S_B$. We estimate a pressure unbalance $\Delta P_{\mathrm{H_2O}}/P_{\mathrm{H_2O}} = (4.6 \pm 0.3) \times 10^{-3}$.
   A discussion on the findings of this chapter is postponed to Sec. 8.2.

6 Force noise decorrelation and performance projection

Identification of contributing effects, time series decorrelation, subtraction, and noise projection for LISA

In the previous chapter, we ascribed the white-noise plateau, and its evolution in time, to the Brownian noise due to the collision of residual gas particles with the TMs. Noise at higher frequencies is already confidently ascribed to interferometer position noise. In this chapter, we focus instead on the sub-mHz frequency band. We have shown in Sec. 4.3 that the detected sub-mHz noise is caused by real forces acting on the TMs, hence it is related to some physical quantities causing such forces.

Our purpose, here, is to understand which sources and mechanisms contributed the most to the measured acceleration noise, and to rule out sources that do not play a relevant role in it.

First, we make a list of some mechanisms which may potentially generate forces on the TMs, and estimate their potential impact on LPF's science measurements. Some of these effects are related to physical quantities monitored onboard LPF. Data series of these quantities are available as telemetries, measured simultaneously to Δg. Some other mechanisms, though, are not associated with any measured time series, and their impact must be estimated differently.

We treat the two cases separately. Wherever time series measurements are available, we apply the decorrelation algorithms described in Chapter 3. This allows us, on the one hand, to estimate correlations and to estimate the impact of the related effects to Δg. On the other hand, it also allows us to estimate the coefficients, "susceptibilities", coupling the time series and acceleration. Such coefficients are often important physical parameters. We analyze a subset of very-long runs, distributed all over the entire mission.
The approach to mechanisms not associated with measured telemetries is completely different. Informed by observations in dedicated experiments, or based on experiments performed on-ground before or after flight, we estimate the contribution to the acceleration noise. Wherever possible, we give the best-estimate values of the contribution. Otherwise, we estimate them in the worst-case scenario.

Finally, we give a final overview summarizing sources contributing to LPF low-frequency noise, and estimate the excess. Results from this chapter will also be presented in [5].

6.1 Physical modelization, assumptions, and methods

The modelization is analogous to what is presented in Chapter 3, and follows closely the nomenclature presented there. First, we note that all the effects that we consider can be effectively linearized, and modeled, as follows.

Let's assume that a time series $y_i(t)$, measured simultaneously[1] to $\Delta g(t)$, is related to an effect generating force on TM(s). The time series $y_i(t)$ are *a priori* affected by noise, likely readout noise, so that they are composed of true signals $z_i(t)$ and noise-only components $n_i(t)$. The differential acceleration, on the other hand, is composed of $\Delta g_0(t)$, unrelated to the considered sources, and the contribution from each source. Each source contributes to TM acceleration through a coupling coefficient, α_i, constant in time and real-valued. In formulas, the model reads

$$\begin{cases} \Delta g(t) &= \Delta g_0(t) + \sum_i \alpha_i z_i(t) \\ y_i(t) &= z_i(t) + n_i(t) \end{cases} \tag{6.1}$$

In the first equation, we state that $\Delta g(t)$ is contributed by all the i signals, in particular by their true-signal components $z_i(t)$. Each signal couples to acceleration through the susceptibility α_i, which is assumed to be constant in time. In addition to the true signals, $\Delta g(t)$ is also contributed by a residual $\Delta g_0(t)$, not related to any of the signals z_i. The meaning of the second equation is that the measured signals $y_i(t)$ are contributed by both the real $z_i(t)$ and the readout noises $n_i(t)$. The model implicitly allows the signals $z_i(t)$ to be cross-correlated. Readout noises, on the contrary, are assumed to be independent, and, obviously, not to contribute to force noise.
The susceptibilities $\boldsymbol{\alpha}$ are real-valued and time-independent numbers, so that their Fourier transform $\boldsymbol{\alpha}(f)$ is real-valued. This has a lot of useful consequences, as analyzed in Chapter 3.

6.1.1 Model identification

The model in Eq. (6.1) is well-representative of all the effects that we take into account. However (see Sec. 3.4), it can not be applied as is. The reason is that this full model, including as unknown variables both the susceptibilities and the readout noises, can not be implemented when decorrelating more than two time series, because of mathematical issues. Some additional simplifications are instead required.
The assumption that we make is that the readout noise affecting the signals is negligible, i.e., that $y_i(t) \sim z_i(t)$.
The model reduces to

$$\Delta g(t) = \Delta g_0(t) + \sum_i \alpha_i y_i(t) \tag{6.2}$$

[1] Our method requires that the time series $y_i(t)$ and $\Delta g(t)$ are measured simultaneously, at the same time stamps. However, if it is not the case, time series can be interpolated at the time stamps of the more frequent series. For example, temperature T was measured with a frequency $f_s = 1/4.8\,\mathrm{Hz}$, whereas acceleration was sampled at 10 Hz, hence we interpolated T at 10 Hz. At the sub-mHz frequencies of interest, the PSD of the interpolated and the non-interpolated time series do not show relevant differences.

The assumption of no noise will be justified in the following. For now, we note that this model is suitable for all the considered effects: thermal gradients, magnetic fields, and imperfect actuation digitization. To some extent, also to fuel depletion, mean temperature, and instrument distortion. In such cases, particularly in the first two, the absence of strong readout noise is observed in measurements: the system included multiple independent sensors showing a correlation high enough to conclude that readout noise is negligible. The case of imperfect digitization is even better, as the corresponding time series are deterministic and have no intrinsic readout error.

In the majority of these cases, the susceptibility α_i is not known, as there is no confident estimate of it. With our analysis, we estimate the residual after decorrelation and the susceptibility.
After applying the model, we analyze to what extent the presence of readout noise might have affected the time series, and the subsequent result.

We analyze separately the first frequency bin, 18 µHz, studying the impact of thermal effects on it. As we note later, temperature effects are the major noise sources affecting very-low frequencies, with a considerable impact. For that analysis, we apply the single-frequency method decorrelating temperature, to better estimate its impact.

6.2 Physical sources, coupling mechanisms

First, we list the available time series known to be coupling with TM acceleration. For each time series, we also describe the coupling mechanism, and, possibly, the expected susceptibility.

Magnetic fields. The interaction with the magnetic fields is another effect to be taken into account. LPF's test masses, because of a residual diamagnetic behavior, interact with the surrounding magnetic fields [72–74]. Without going into the details, different strategies were implemented to reduce this effect as much as possible, such as reducing the TM magnetic susceptibility and the magnetic moment, and reducing the onboard-generated magnetic fields and gradients. TMs generally interact with magnetic fields through the Lorentz force formula

$$\boldsymbol{F} = \int_V \left(\boldsymbol{M}\cdot\boldsymbol{\nabla}\boldsymbol{B} + \boldsymbol{J}\times\boldsymbol{B}\right) dV. \tag{6.3}$$

The first term contains magnetization terms, whereas the second term describes the effect of currents within the TM bodies. The magnetic susceptibility of the TMs has been measured $\chi \sim -2.5\times10^{-5}$, and the magnetic moment is zero within instrument error $4\,\mathrm{nA\,m^2}$ [15]. A few considerations lead us to a usable expression of the magnetic force:

1. Accounting for all the potential magnetic interaction mechanisms, and estimating their magnitude [72], it is possible to conclude that the most relevant contribution comes from the induced magnetization. Force on a single test mass, along x, reads
$$F_x = \int_V \boldsymbol{M}\cdot\partial_x\boldsymbol{B}\,dV = \frac{\chi}{\mu_0}\int_V \boldsymbol{B}(t)\cdot\partial_x\boldsymbol{B}(t)\,dV, \tag{6.4}$$

where μ_0 is the vacuum magnetic permeability, and $\partial_x \boldsymbol{B}$ the magnetic gradient along the x direction. In the last step, we have fully introduced the time dependence.

2. Magnetic field and its gradient along x have been measured onboard, with four tri-axial magnetometers located in the spacecraft, on the xy plane, about 15-20 cm from the TMs. The four magnetometers show consistent results, highly correlated below $\sim$ mHz frequencies. Moreover, the measured magnetic fields are compatible with measurements of other missions located in L1, such as NASA's DSCOVR. This shows that $B(t)$ is confidently dominated by the interplanetary magnetic field. The magnetic gradient, on the contrary, is dominated by local sources, as the large-scale interplanetary magnetic gradient is negligible.

3. Linearizing the previous force expression to the first order, and expressing it as differential acceleration, it reads

$$\Delta g(t) = \frac{\chi L^3}{M\mu_0}(\partial_x \boldsymbol{B}_{\mathrm{DC},2} - \partial_x \boldsymbol{B}_{\mathrm{DC},1}) \cdot \boldsymbol{B}(t). \tag{6.5}$$

L and M are the TM side length and mass, $\boldsymbol{B}(t)$ is the average measured magnetic field, $\partial_x \boldsymbol{B}_{\mathrm{DC},i}$ is the average static magnetic gradient at the location of TMi, averaged over the TM x-surfaces.

Equation (6.5) describes the linearized interaction of the (3-D) magnetic field with Δg, and defines the corresponding susceptibilities α_{B_x}, α_{B_y}, α_{B_z}:

$$\alpha_{B_\zeta} = \frac{\chi L^3}{M\mu_0}(\partial_x \boldsymbol{B}_{\zeta,\mathrm{DC},2} - \partial_x \boldsymbol{B}_{\zeta,\mathrm{DC},1}), \quad \zeta = \{x, y, z\} \tag{6.6}$$

Nevertheless, there is no reliable measurement of the susceptibilities, as there is no precise measurement of the magnetic gradient at the TM locations. Despite magnetic gradients of the order of $600\,\mathrm{nT\,m^{-1}}$ have been measured at the magnetometers' location, gradients at the TM location might be different because of the presence of magnetic thermistors on the EHs. Magnetometers were located too far away from the TMs to properly measure the local magnetic gradients. Reference [72], however, estimates that magnetic gradients at TM location of the order of $10\,\mu\mathrm{T\,m^{-1}}$ could be plausible.

The time series $B_x(t)$, $B_y(t)$, and $B_z(t)$ are built averaging the measurements of the four magnetometers. Regarding the impact of readout noise on the average magnetic field, the single-measurement time series show a high correlation below mHz frequencies, suggesting that the readout noise is negligible.

Mean temperature and thermal gradients. Temperature is known to be related to a number of effects inducing stray forces on the TMs. Both the average temperature and its spatial gradient [44, 71] are indeed related to force-inducing effects.
Thermal effects are primarily associated with the outgassing from the internal surface of the VEs. As widely described in the previous chapter, outgassing decreased over the course of the mission, hence the incidence of thermal-induced forces is expected to

decrease in time. Indeed, we will note that it was at its maximum during the first months of the mission, and decreased thereafter.

Within each GRS, eight thermistors were located on the outer EH faces, four on the $x-$ faces, and four on $x+$, with a sensitivity of about $10\,\mathrm{mK\,Hz^{-1/2}}$ at 20 µHz. Also, differential measurements were available, performed with Wheatstone bridges among sensors on opposite faces. The differential channels allowed for improved sensitivity to temperature gradients, down to $0.1\,\mathrm{mK\,Hz^{-1/2}}$ at 20 µHz. We also find that the differential measurements are sufficiently sensitive across the entire sub-mHz band. Single thermometers, on the other hand, lose coherence above 30 µHz, where the electronic noise becomes limiting. However, its PSD is so steep ($\propto f^{-7.2}$ [44], which also includes the PSD of the readout noise, much steeper than the f^{-2} low-frequency branch of Δg) that its impact on frequencies > 36 µHz is generally negligible. It follows that, if on the one hand, it is true that noise at 18 µHz is explainable with thermal effects, it becomes quickly negligible at higher frequencies.

Averaging measurements from the sensors, we build the mean temperature time series $\overline{T}(t)$, and the two time series of the differential measurements, $\Delta T_1(t)$ and $\Delta T_2(t)$, respectively referring to GRS1 and GRS2. The contribution of temperature to Δg eventually reads

$$\begin{cases} \Delta g_{T\ \mathrm{mean}}(t) &= \alpha_T\, \overline{T}(t) \\ \Delta g_{T\ \mathrm{grad}}(t) &= \alpha_{\Delta T_1}\, \Delta T_1(t) + \alpha_{\Delta T_2}\, \Delta T_2(t) \end{cases} \tag{6.7}$$

The susceptibilities are not precisely known, only estimations are available. After decorrelating, we will however compare the resulting α_T with the one from the linear fit in Chapter 5. Given the nature of the time series and the impact of readout noise on $T(t)$, we include thermal effects as follows:

- We consider thermal gradients among all the other time series in the multi-frequency setup, as their noise is negligible in the analyzed frequency band. We expect to find susceptibilities of the order of $10\,\mathrm{pm\,s^{-2}\,Hz^{-1/2}}$. We implicitly assume that the time delay between thermal gradients and the induced forces is negligible: indeed we find that the role of thermal gradients is negligible, hence a phase-shift would not make a relevant difference.

- We also include the mean temperature in the multi-frequency analysis. Even though it is affected by readout noise, we already analyzed that the application of the no-noise model to cases with readout noise can not lead to overestimation of the contribution or the susceptibilities (Sec. 3.6.2): if the contribution itself is negligible, the presence of noise will also lead to a negligible estimate. The impact of the mean temperature is expected to be dominant at the first frequency.

- We analyze separately the first frequency, 18 µHz, studying the impact of thermal effects on it. Indeed, thermal effects are dominant at lower frequencies, making it meaningful to treat them separately. We compare susceptibilities with the ones we find in the DC analysis.

Imperfect digitization of actuation waveforms. This effect is a direct consequence of LPF's imperfect actuation digitization [47]. LPF's actuation digitization scheme was affected by a slight round-off error, affecting all the 24 GRS electrodes, and altering the quasi-sinusoidal actuation waves. This resulted in the application of unwanted low-frequency voltages $\delta V_i(t)$ on the electrodes, of the order of $\sim$ µV. The effect has been identified and characterized, and the time series $\delta V_i(t)$ have been produced with reverse-engineering methods. Since $\delta V_i(t)$ is deterministic, there is no associated error.

Quasi-static voltages generate force, coupling to the TM voltage and hence to the TM charge, which is linearly varying in time. The TM charge follows $q(t) = q_0 + \dot{q}t$. A possible model, including both constant and linear terms, would read

$$\Delta g(t) = \sum_{i\in[1,24]} a_i\, \delta V_i(t) + \sum_{i\in[1,24]} b_i\, t\, \delta V_i(t) \tag{6.8}$$

To reduce the number of free parameters, we introduce a few simplifications. First, we only consider x electrodes, neglecting the y and z ones[2]. Second, in the time-varying part, we only consider its contribution through the effective potentials $\delta\Delta_{x,j}(t)$. These are defined, as usual, by combinations of the voltages on the four x electrodes on the x faces: $\delta\Delta_{x,j} = (\delta V_{\mathrm{x1},j} + \delta V_{\mathrm{x2},j} - \delta V_{\mathrm{x3},j} - \delta V_{\mathrm{x4},j})$, where j is an index indicating the corresponding GRS. The model reads

$$\Delta g(t) = \sum_{i\in[1,8]} \alpha_i\, \delta V_{x,i}(t) + \sum_{j\in 1,2} \beta_j\, t\, \delta\Delta_{x,j}(t) \tag{6.9}$$

The susceptibilities α_i are not known. The susceptibilities β_j can however be forecast, since they are related to the charging rate by

$$\beta_j = (-)^{j+1}\, \frac{\dot{q}_j}{C_T} \left|\frac{\partial C_x}{\partial x}\right| \tag{6.10}$$

We will make a comparison in the end between $\dot{q}_j$ from our fit and from measurements. As the posteriors for $\dot{q}_1$ and $\dot{q}_2$ are expected to be similar, as measured in charge experiments [26] and discussed in Sec. 6.3, we make a further simplification of the model, unifying $\dot{q}_1$ and $\dot{q}_2$ as

$$\Delta g(t) = \sum_{i\in[1,8]} \alpha_i\, \delta V_{x,i}(t) + \beta\, t\, (\delta\Delta_{x,2}(t) - \delta\Delta_{x,1}(t)) \tag{6.11}$$

However, we separately check that the posteriors of $\dot{q}_1$ and $\dot{q}_2$ are really compatible. With these simplifications, imperfect digitization is modeled with 9 time series.

Instrument distortion. This is the well-known mechanism already described in Sec. 2.2.6, which induces gravitational force because of the µm deformation of the

[2]However, we also estimated the contributions of the y and z in a previous version of this analysis, finding that it is not relevant. Rather, including them would require a high number of parameters, which is not desirable.

GRS structure. We recall that the deformation induces Δg noise through

$$\Delta g = -\omega_d^2(\Delta x_{\mathrm{GRS}} - \Delta x_{\mathrm{OMS}}). \tag{6.12}$$

The stiffness, ω_d^2, is known to be $-3.32 \times 10^{-7}\,\mathrm{s}^{-2}$ with 1.5% accuracy, hence it is practically known exactly. Additional readout noise is, on the other hand, not completely known.

The readout noise of this time series is dominated by the electronic noise in Δx_{GRS}, needing more attention. When we first applied the estimation method with the time series $(\Delta x_{\mathrm{GRS}} - \Delta x_{\mathrm{OMS}}) \sim \Delta x_{\mathrm{GRS}}$, we got a result showing a high level of correlation between Δx_{GRS} and Δg, peaking at about 0.1-0.3 mHz. Such a correlation might hint at a high impact of instrument distortion on TM acceleration. We note that, fortunately, that is not the case: a correction, which we describe here, leads to different results. In Fig. 6.1, we show the ASD of Δx_{GRS}, before and after that correction.

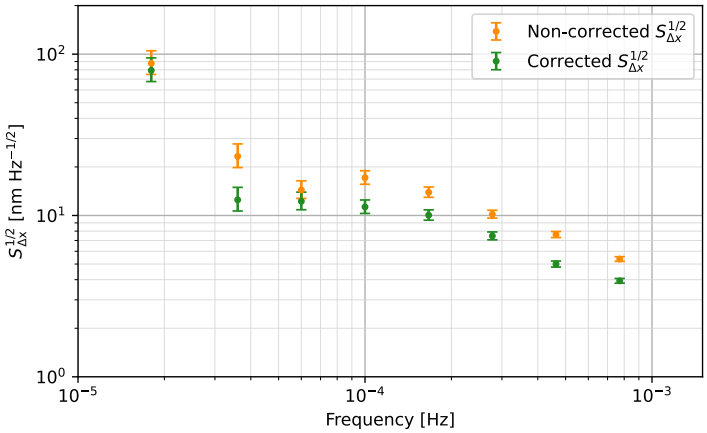


Figure 6.1: Estimated PSD of instrument distortion Δx_{GRS}, run Feb17B. This estimate is valid only at very low frequencies, as it is presumably affected by electronic readout noise in the LISA Band. We observe how the correction presented in the text is effective in correcting the spurious noise correlation to Δg.

The time series Δx_{GRS} has been obtained as a linear combination of four measurements from four electrode pairs, two for each TM, $\mathrm{x1}_i$ and $\mathrm{x2}_i$. For the i-th GRS, the position x_i is obtained as $x_i = (\mathrm{x1}_i + \mathrm{x2}_i)/2$. The same capacitive bridges are used to measure rotations, as $\phi_i = (\mathrm{x1}_i - \mathrm{x2}_i)/(2r_0)$, where r_0 is an effective arm of 11.15 mm. Measurements from $\mathrm{x}j_1$ and $\mathrm{x}j_2$ were handled by different boards. We note that Δx_{GRS} could be measured, equivalently, by

$$\Delta x_{\mathrm{GRS}} \sim \frac{(\mathrm{x21} + \mathrm{x22})}{2} - \frac{(\mathrm{x11} + \mathrm{x12})}{2} \sim (\mathrm{x21} - \mathrm{x11}) \sim (\mathrm{x22} - \mathrm{x12}). \tag{6.13}$$

The last two combinations provide two ways to evaluate Δx_{GRS} which do not use all four bridge measurements. In principle, the three combinations should provide equivalent measurements, with the advantage of the first one to reject angular motion. Nevertheless, we note that the third combination is affected by a strong spurious cross-noise with Δg, which is likely cross-talk between sensing and actuation, hence

not representing any GRS displacement. Therefore, we use a two-bridge corrected combination for $\Delta x_{\rm GRS}$:

$$\Delta x_{\rm GRS} \sim \Delta x_{\rm GRS} + r_0 \Delta\phi_{\rm GRS} \sim (\mathrm{x21} - \mathrm{x11}). \tag{6.14}$$

For coherence, we also subtract $(\Delta x_{\rm OMS} + r_0\Delta\phi_{\rm OMS})$, even though it does not make any relevant difference.
We include this time series in the analysis, but we analyze its readout noise later, see point 4 on page 118.

Fuel depletion. Gravitational force from fuel depletion might induce in-band noise, if the mass loss spectrum has some components in the science frequency band. This is the same mechanism described in Sec. 2.1.4. In Sec. 5.3 we showed that it contributes the most at zero-frequency, as a quasi-static DC drift. The induced acceleration $\Delta g_F(t)$ has already been computed from the fuel mass loss time series, hence the expected susceptibility is known to be $\kappa_t\kappa_b \sim 1$ within 10%. Here, the κs are the calibration coefficients described in the previous chapter. The time series $\Delta g_F(t)$ is well known, but its additional noise, related to the flowmeter noise, is unknown at LISA-band frequencies. We include this time series in the analysis, but we analyze its readout noise later, see point 4 on page 118.

6.3 Multi-frequency decorrelation

Among the time series that we include in the multi-frequency decorrelation set, we list: the average magnetic field vector (3 time series), thermal gradients (2 time series), mean temperature (1 time series) imperfect digitization time series (9 time series), fuel mass depletion (1 time series), instrument distortion (1 time series). We apply the multi-frequency decorrelation method described in Sec. 3.6, decorrelating these 17 time series in the frequency band [18 µHz–0.77 mHz].
Before proceeding, we need to highlight a few points:

1. It is essential that the number of available periodograms exceeds the number of decorrelated signals, to make the result significant. Moreover, if one wants the Wishart distribution to be satisfied at single frequencies, it is necessary that the number of periodograms available at each frequency exceeds the number of decorrelated time series. Moreover, including frequencies above mHz is not necessary, as they are dominated by Brownian noise.

   Given these observations, we restrict the application of the decorrelation to the longest runs, > 8.6 d – Apr16A, Nov16B, Dec16B, Feb17B, Jun17. We identify as Jun17 the union of Jun17A and Jun17B, which is possible as the two runs were performed in similar experimental conditions, and showed a similar noise. In run Feb17B, the frequency band [18 µHz–0.1 mHz] contains 61 periodograms, and the entire [18 µHz–0.77 mHz] contains 538. The shortest run of this set, Nov16B, contains 38 periodograms below 0.1 mHz.

2. We apply the decorrelation algorithm, with an MCMC method estimating the posteriors for the residual noise and real-valued susceptibilities. We use a slightly modified version of the likelihood in Eq. (3.49): we parametrize the residuals, $S_{\Delta g_{0,f}}$, as being composed of a frequency-independent Brownian noise S_B and a frequency-dependent excess $S_{E,f}$, so that

$$S_{\Delta g_{0,f}} = S_B + S_{E,f} \tag{6.15}$$

This parametrization allows us to properly estimate the residuals, intended as residuals after decorrelation of the selected time series and the Brownian noise. As a prior for the Brownian S_B, we employ the posterior that we found in the Brownian evaluation analysis, Sec. 4.1. We employ a log-flat prior for the residual PSDs and the Brownian PSDs, and flat priors for the αs.
Overall, the model reads (see Eq. (3.49)):

$$\begin{aligned} p(\boldsymbol{\alpha}, S_B, \{S_E\}) &\propto \\ &\propto \prod_f \frac{1}{S_{\Delta g_{0,f}}^{M_f}} \exp\left[-\frac{\boldsymbol{\eta}\cdot\boldsymbol{W}_f\cdot\boldsymbol{\eta}^T}{S_{\Delta g_{0,f}}}\right] \times \mathrm{pr}(S_B) \times \frac{1}{S_B}\frac{1}{S_{E,f}}, \end{aligned} \tag{6.16}$$

where $S_{\Delta g_{0,f}} = S_B + S_{E,f}$ and $\boldsymbol{\eta} = (-1, \boldsymbol{\alpha})$.

3. Decorrelating, we find an estimate of the residuals and the susceptibilities. In addition to these, we estimate the contribution of each block of sources as described in Eqs. (3.51) and (3.52). Namely, blocks are: the 3 time series of the magnetic fields; the mean temperature; the 2 time series of thermal gradients; the instrument distortion; the fuel depletion; the 9 time series of imperfect digitization.

We proceed as follows:

1. First, we apply the decorrelation algorithm to eight frequencies, in the band [18 µHz – 0.77 mHz], decorrelating all 17 time series previously described. We employ the presented procedure, even including those frequencies whose number of periodograms is lower than the number of decorrelated time series.

2. We apply the decorrelation algorithm to all runs, decorrelating all 17 time series, including only those frequencies with a number of periodograms higher than the number of decorrelated time series, hence satisfying the Wishart condition $M_f \geq p$. This means that we perform the analysis starting from 100 µHz, except for Dec16B, which is longer and allows us to start from 60 µHz. This is intended as a test of the validity of the previous analysis; we find compatible results.

3. We apply the decorrelation algorithm to the longer runs, Dec16B, Feb17B, Jun17, decorrelating only the disturbance time series (8 time series). These runs have at least 9 periodograms at the lowest frequency, hence always satisfying the Wishart condition on the entire frequency band.
Again, this is intended as a test of the validity of the first analysis, and we find compatible results.

4. At last, we apply the decorrelation algorithm to run Dec16B, decorrelating only the imperfect digitization time series. This is the only run for which it is possible to have $M_f \geq p$ over the entire frequency band, hence we can use this run to check the validity of the decorrelation of imperfect digitization in the first case. Again, we find compatibility.

6.3.1 Multi-frequency decorrelation below 1 mHz

As always, we take as a representative example run Feb17B, which we show in Fig. 6.2. We show the result of the multi-frequency decorrelation, the measured $S_{\Delta g}$ excess over Brownian. Moreover, we estimate the contribution of the individual blocks of time series, wherever possible (Eq. (3.52)), and the total contribution from the considered sources, wherever all individual contributions can be calculated. We note that the residuals after decorrelation (red points) almost completely overlap with the original signal, confirming that these sources have a negligible contribution to $S_{\Delta g}$, in this run. The highest contribution fraction is found at 18 µHz, and, as we noted previously, is related to a (relatively) high correlation with the mean temperature. We postpone this discussion to Sec. 6.4, where the first frequency bin is discussed in more detail.

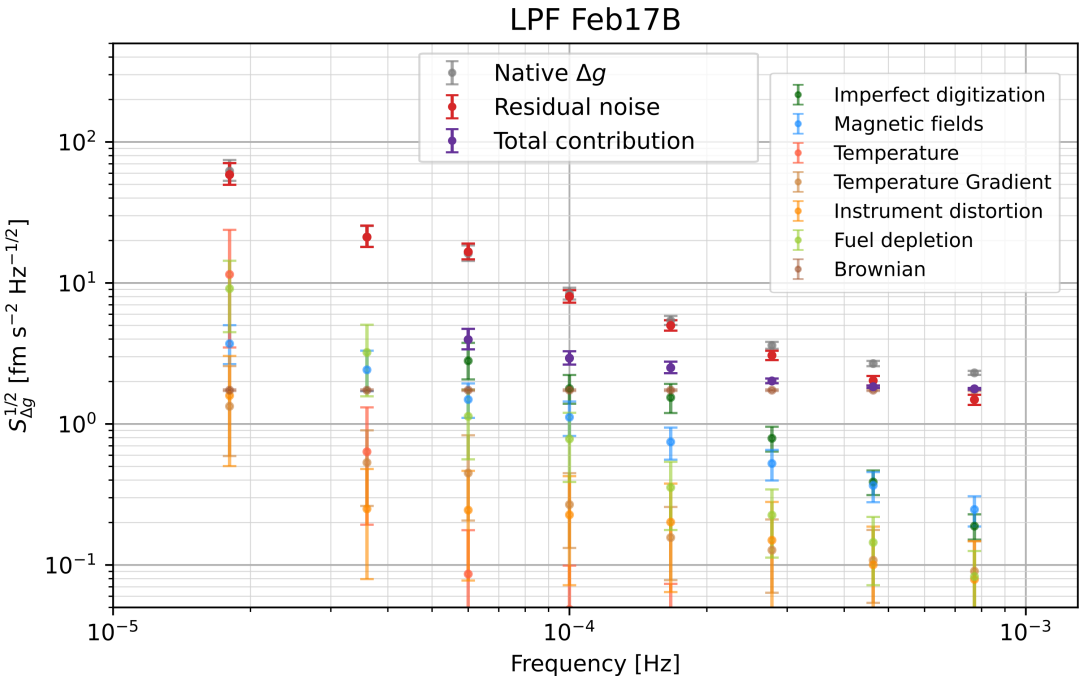


Figure 6.2: Results of multi-frequency time series decorrelation of the reference run Feb17B, in the frequency band [18 µHz–0.77 mHz]. Grey points represent native data, and red points represent the residual after decorrelation. Light points represent the individual contribution of each source, shown wherever possible to calculate. Dark blue points represent the estimates of the total contribution from the considered sources, shown wherever all individual contributions were available. We note that, in run Feb17B, the major contribution is found at 18 µHz, and is related to the correlation to mean temperature.

In Fig. 6.3, we provide the posterior susceptibilities, as histograms of the posterior samples from the MCMC chains. The top left panel shows the magnetic susceptibilities α_{B_ζ}, also expressed as average magnetic gradients through $\partial_x B_\zeta / \alpha_{B_\zeta} = (M\mu_0)/(\chi L^3)$. The top central panel shows the susceptibility α_T, and the top right panel shows the thermal gradients susceptibilities $\alpha_{\Delta T_i}$. In the bottom panel, we show the mean charging rate $\dot{q}$, computed from β as in Eq. (6.10).

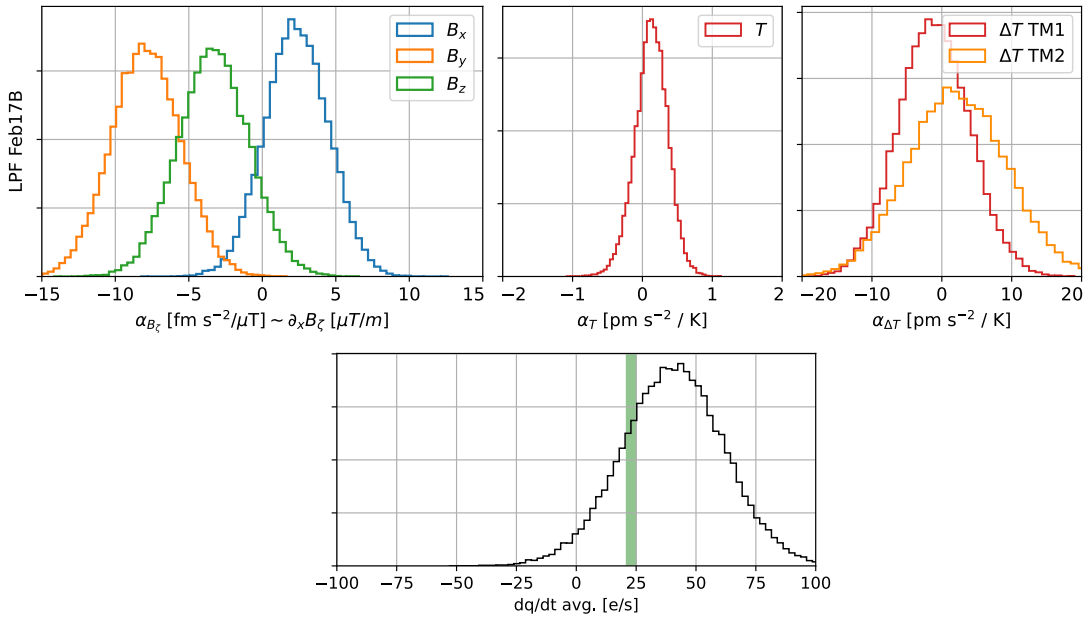


Figure 6.3: Posterior susceptibilities from the MCMC chains of the time series decorrelation, run Feb17B, probability distributions [a.u.]. The top left panel shows susceptibilities of the magnetic fields, proportional to the average magnetic field gradients. The top central and the right panel show, respectively, the susceptibilities of the average temperature and the thermal gradients. The bottom panel shows the mean charging rate $\dot{q}$, calculated following Eq. (6.10).

We make a few observations:

1. Susceptibilities α_{B_x} and α_{B_z} are compatible with zero, within $2\sigma \sim 4\,\mathrm{fm\,s^{-2}\,\mu T^{-1}}$. On the contrary, α_{B_y} is not: its estimate is $\alpha_{B_y} = (-7.9 \pm 2.4)\,\mathrm{fm\,s^{-2}\,\mu T^{-1}}$, which is significantly different from zero at more than 3σ. This susceptibility corresponds to an average magnetic field gradient $\partial_x B_y = (-7.9 \pm 2.4)\,\mathrm{\mu T\,m^{-1}}$, which is reasonably compatible with the expected gradient generated by the thermistors on the EH (of the order of $10\,\mathrm{\mu T\,m^{-1}}$ [72]).

2. The susceptibility α_T is compatible with zero, but it is also compatible with the value measured from the quasi-DC evaluations, within 1σ (see Fig. 5.4).
The susceptibility of thermal gradients is compatible with the expected order-of-magnitude estimate.

3. The average charging rate is compatible with expectations. In addition to the presented estimations, we also estimated the single TM charging rates, without averaging $\delta\Delta_x$ in Eq. (6.11). The estimated rates, $\dot{q}_1 = (19 \pm 14)\,\mathrm{e\,s^{-1}}$ and $\dot{q}_1 = (24 \pm 14)\,\mathrm{e\,s^{-1}}$, are consistent with the measurements from charging

experiments, respectively (22.9 ± 2) and $(24.5 \pm 2)\,\mathrm{e\,s^{-1}}$, though with a larger error bar.

4. **Readout noises: fuel and distortion.** Noisy time series, i.e. fuel depletion, instrument distortion, and temperature, need further explanation. In Sec. 3.6.2, we studied the effect of the presence of readout noise on the decorrelation efficiency and parameter estimation. The main result was that, if a time series is affected by readout noise, the estimated susceptibility and the estimated contribution drop proportionally to the noise fraction. Indeed, this is exactly what we find for fuel depletion and instrument distortion. For instance, in the case of fuel depletion we find α_F peaking at about 0.5, instead of the expected 0.92. This would mean that a fraction $\sim 40\%$ of the measured PSD is due to readout noise. The same is true for instrument distortion, which yields an estimated ω_d^2, which is less than 1/3 of the known one.
Following this observation, we find that the estimation of the contribution of fuel depletion and instrument distortion in Fig. 6.2 is biased, giving results systematically lower than the real ones.

   A more realistic estimation of the contribution of these two sources with noisy readouts can be carried out with some tools developed in Sec. 3.6.2. In particular, let's focus on the single-frequency case, and let's suppose that we need to decorrelate a signal S_{zz}. The problem of estimating the contribution is that *a)* we do not have access to the true signal S_{zz}, but rather to its noise-corrupted version, $S_{yy} = (S_{zz} + S_{nn})$, and *b)* we do not have access to the true susceptibility α, but rather to its corrupted version $\alpha^{(*)}$. We proved that the estimation of α in case of noise leads to the noise-corrupted version

$$\alpha^{(*)} = \alpha \, \frac{S_{zz}}{S_{zz} + S_{nn}} \tag{6.17}$$

   We use the last result to our advantage. Indeed, for these two cases we do have access to the true susceptibilities, which we know from calculations.
We can rewrite the true contribution as follows:

$$S_{\mathrm{contr}} = \alpha^2 \, S_{zz} = \alpha \, \alpha S_{zz} = \alpha \, \alpha^{(*)} \, \frac{S_{zz} + S_{nn}}{S_{zz}} S_{zz} = \alpha \, \alpha^{(*)} \, S_{yy} \tag{6.18}$$

   According to this equation, the errors in the estimation of α and S_{yy} compensate, so that in the end S_{contr} is a good estimate of the noise contribution. This means that a valid estimate of the true contribution to noise can be obtained using the noisy-corrupted estimations $\alpha^{(*)}$, and S_{yy}, provided that the true susceptibility α is reasonably known *a priori*. In our case, it is exactly what happens for the instrument distortion and fuel depletion time series. However, this only applies in the single-frequency case.

   In Fig. 6.4, we make a comparison, for the single-frequency case in run Feb17B, comparing the straightforward calculation, careless of the presence of readout noise, and the result in Eq. (6.18). We estimate $\alpha^{(*)}$ and S_{yy} with numerical tools, using the usual likelihood in Eq. (3.49). The light points represent the

ideal-case calculation, careless of the presence of readout noise, and the darker ones represent our better estimate.

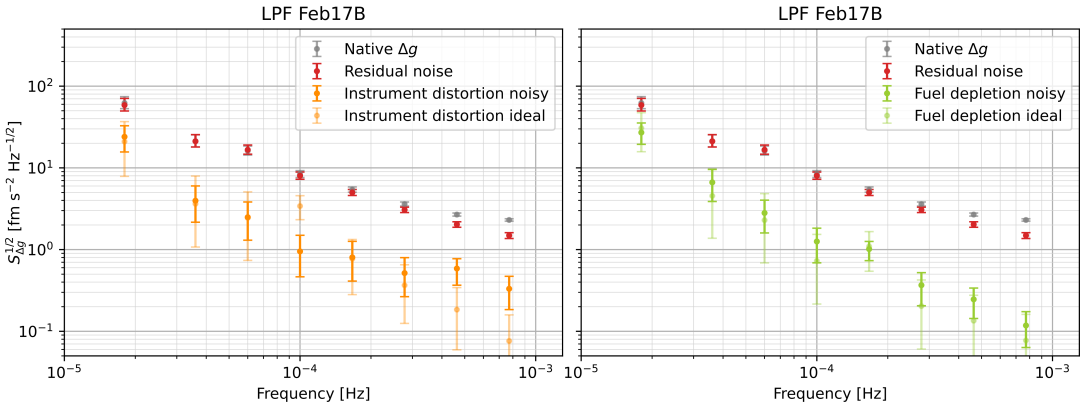


Figure 6.4: Estimated contribution from noisy time series Δx_{GRS} and Δg_F (left panel and right panel, respectively), run Feb17B. The light points represent the calculation careless of the presence of readout noise, while the darker ones represent our better estimate, following Eq. (6.18). Red and grey points have the same meaning as in Fig. 6.2.

First, looking at the contribution of fuel depletion, we note that the better estimate is indeed higher than the previous one, as expected. The resulting estimate of the noise contribution is lower than the observed excess, and the conclusion is the same as with the previous analysis.

Instrument distortion is a bit different. Except for a feature at 0.1 mHz, the overall contribution is lower than 10% power at all frequencies $\geq 36\,\mu$Hz, and can not explain the $1/f$ dependence of $S^{1/2}_{\Delta g}(f)$. We observe that the feature at 0.1 mHz is due to a negative susceptibility at that particular frequency, causing incompatibility with our noise model. We think that that feature might have the same spurious cross-talk origin observed in the other Δx capacitive channel: in that case, the latter model (dark orange) would apply for the estimation. If this were not the case, and that was effectively a force disturbance, then the ideal-case model (light orange) would apply: in that case, this contribution would however account for no more than 25% power, localized at 0.1 mHz.

In the following, analyzing other LPF runs, we include Δx_{GRS} and Δg_F as we did previously in Fig. 6.4.

5. **Readout noises: temperature.** Temperature falls into this category too, for frequencies higher than 36 µHz. The main difference with the previous two cases is, however, that noise is irrelevant at 18 µHz, which effectively constrains the susceptibility. This means that the estimated contribution at higher frequencies can be intended as an upper limit, mainly contributed by noise. Since its effect is already negligible, we conclude that temperature is negligible, even considering the presence of readout noise.

Other long runs – decorrelation and susceptibilities

In the next few pages, we show the results of the time series decorrelation of the long runs Apr16A, Nov16B, Dec16B, and Jun17. The conclusions are globally unchanged:

1. The magnetic susceptibility α_{B_x} is always compatible with zero. As in run Feb17B, α_{B_y} is globally not compatible with zero, and a slight deviation from zero of α_{B_z} is also noted in run Nov16B. The measured values are consistent with the scenario that thermistors are the major contributors to DC magnetic field gradient at TM locations, with magnetic gradients of the order of $10\,\mu\mathrm{T\,m^{-1}}$.

2. The thermal susceptibility α_T is compatible with the quasi-DC linear fit, as in Fig. 5.4. However, uncertainties are always larger than those from the linear fit.

3. The average charging rate, calculated from the susceptibility of the imperfect-digitization time series, is compatible with results from charging experiments.

4. Temperature always has a large contribution to the measured PSD in the first frequency bin. This is discussed further in the next section.

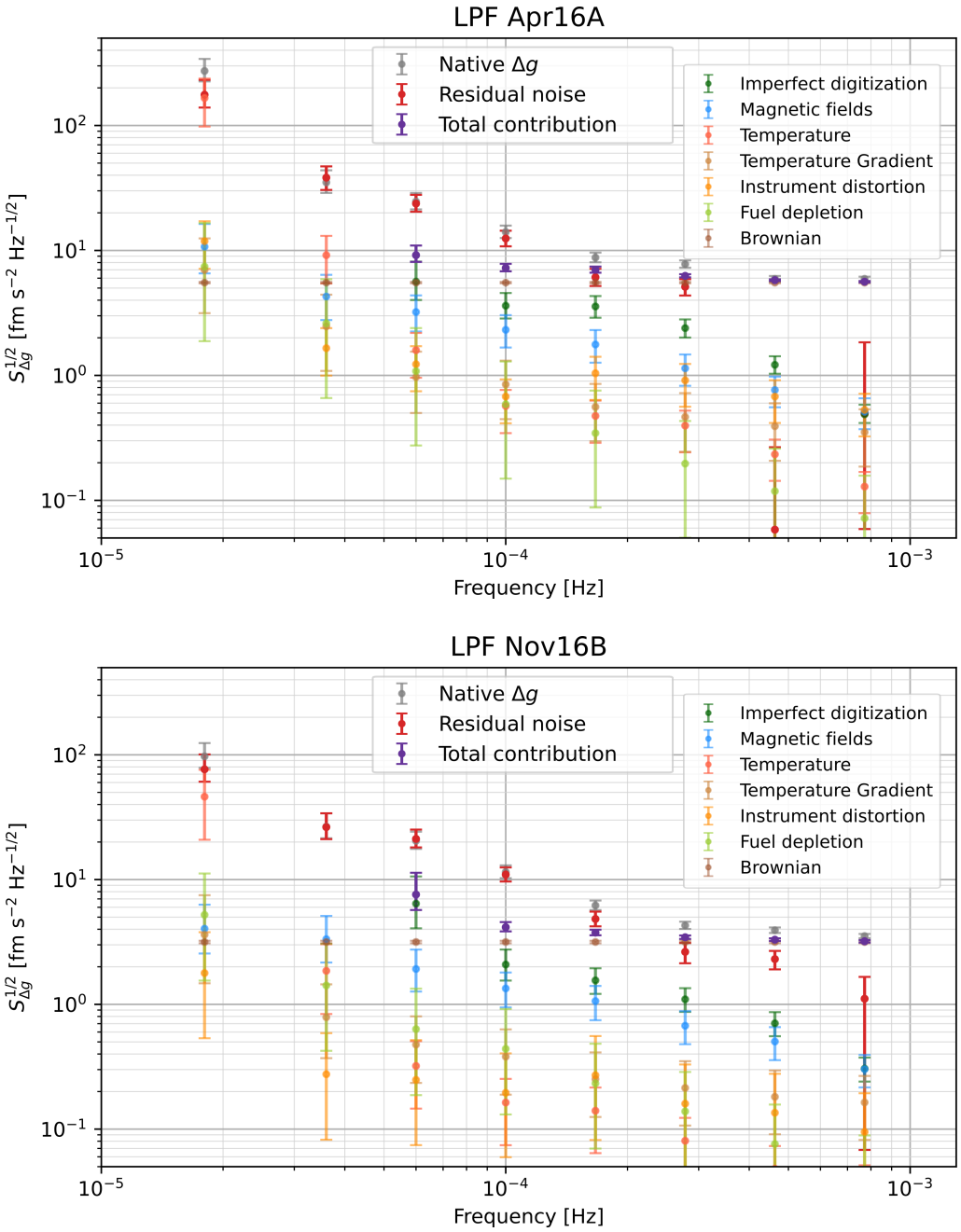


Figure 6.5: Results of multi-frequency time series decorrelation of runs Apr16A and Nov16B, in the frequency band [18 µHz–0.77 mHz]. Grey points represent native data, and red points represent the residual after decorrelation. Light points represent the individual contribution of each source, shown wherever possible to calculate. Dark blue points represent the estimates of the total contribution from the considered sources, shown wherever all individual contributions were available.

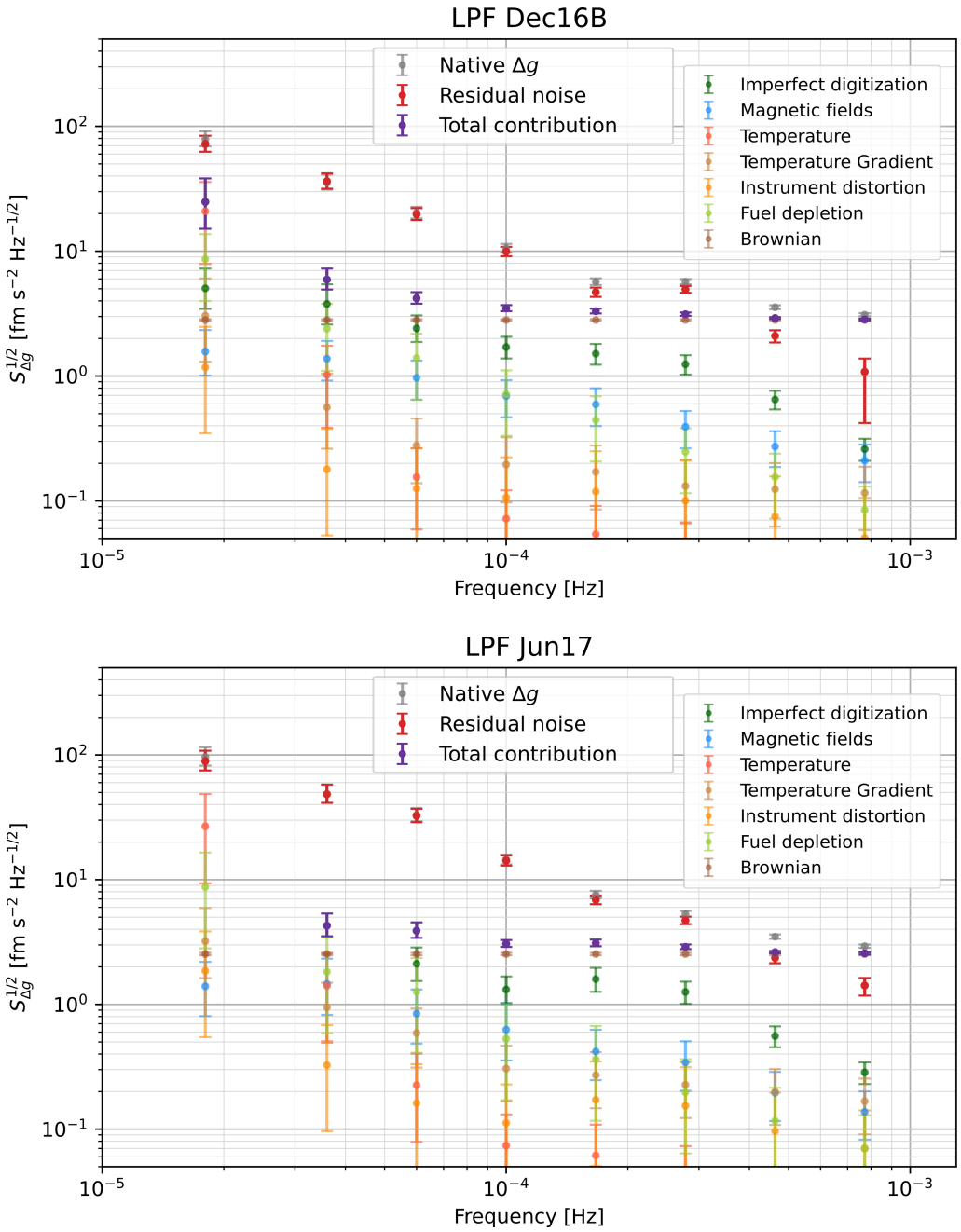


Figure 6.6: Results of multi-frequency time series decorrelation of runs Dec16B and Jun17, in the frequency band [18 µHz–0.77 mHz]. Grey points represent native data, and red points represent the residual after decorrelation. Light points represent the individual contribution of each source, shown wherever possible to calculate. Dark blue points represent the estimates of the total contribution from the considered sources, shown wherever all individual contributions were available.

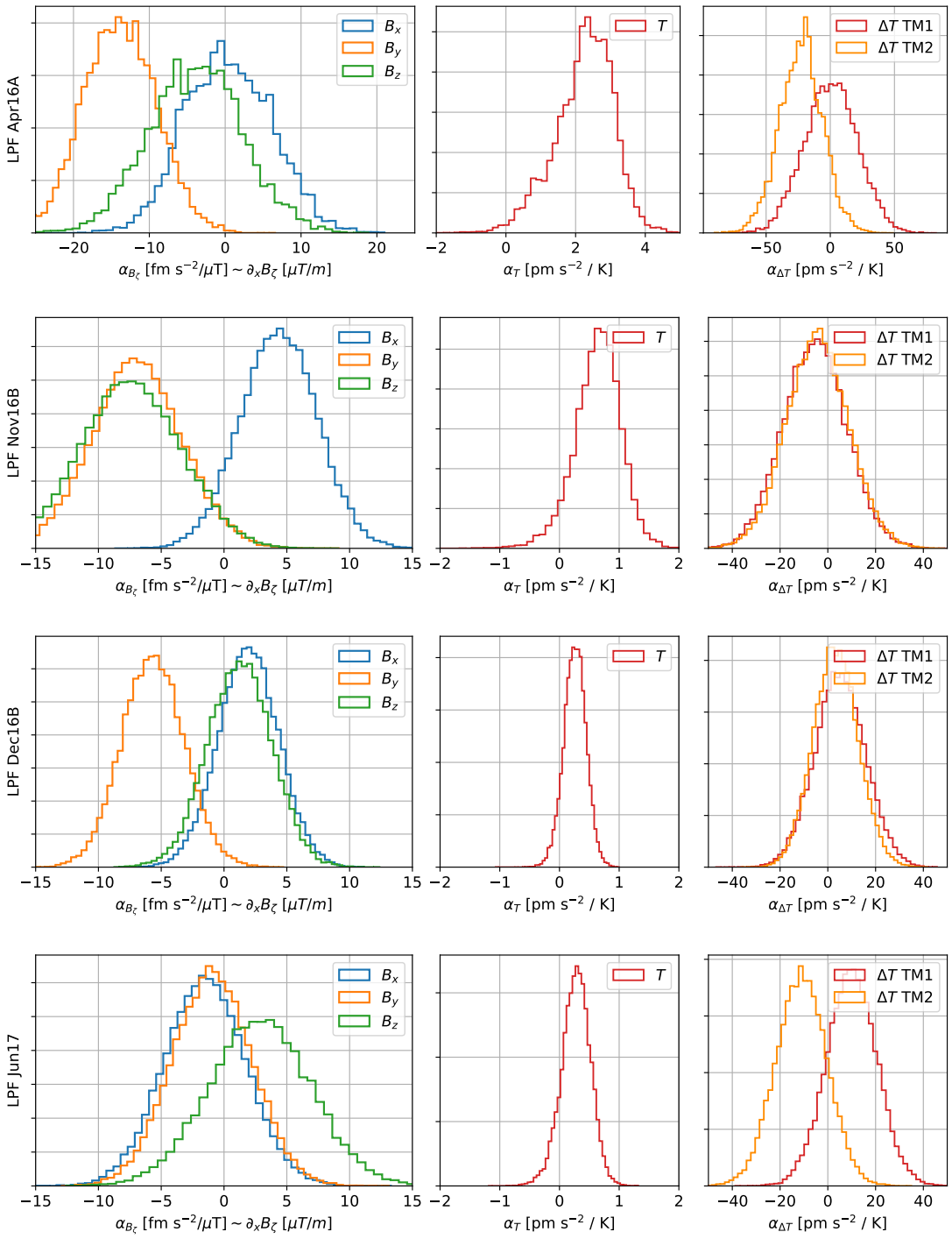


Figure 6.7: Posterior susceptibilities from the MCMC chains of the time series decorrelation, runs Apr16A, Nov16B, Dec16B, Jun17. Probability distributions [a.u.]: magnetic fields, average temperature, and thermal gradients.

6.3.2 Decorrelation – Wishart distribution check

We provide, for comparison, the results of the decorrelation of the full set of 17 time series, only in the cases for which $M_f \geq p$ for all frequencies f. We provide results for run Dec16B, starting from 60 µHz, and Feb17B, starting from 100 µHz. We also analyzed runs Apr16A, Nov16B, and Jun17, always finding full compatibility with results in the complete case.

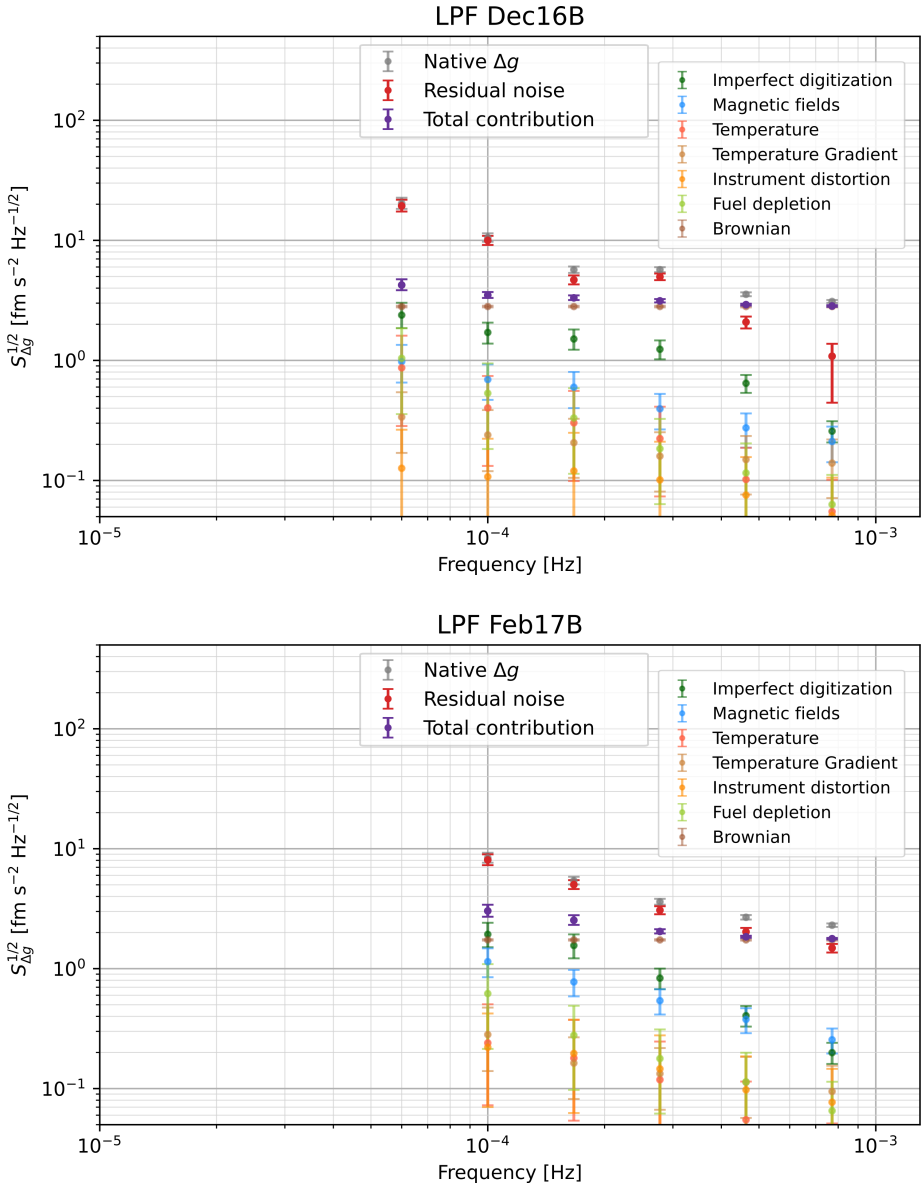


Figure 6.8: Check of the validity of the full-band result, runs Dec16B and Feb17B. We decorrelate the entire set of 17 time series, only on frequencies for which the number of periodograms is greater than 17. We find full compatibility with results in the complete case.

6.3.3 Decorrelation of non-electric disturbances

We provide, for comparison, the results of the decorrelation of the set of 8 time series, only including non-electric disturbances (magnetic fields, temperature, thermal gradients, fuel depletion, and instrument distortion), excluding imperfect digitization time series. This way, the condition $M_f \geq p$ applies for all frequencies f, in runs Dec16B, Feb17B, Jun17. We provide results for run Dec16B and Feb17B. We find full compatibility with results in the complete case.

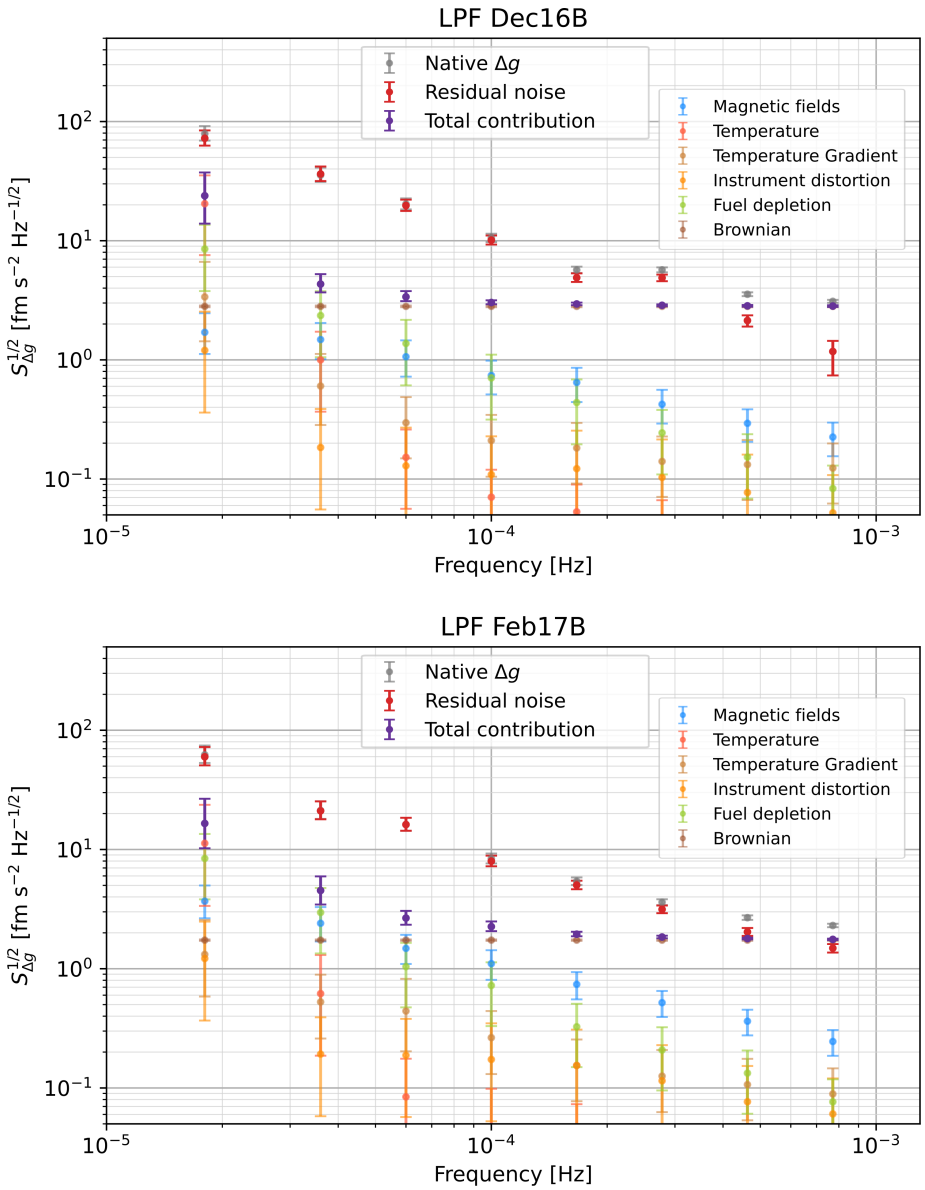


Figure 6.9: Check of the validity of the full-band result, runs Dec16B and Feb17B. We decorrelate the subset of 8 time series, only including non-electric disturbances, excluding imperfect digitization time series. We find full compatibility with results in the complete case.

6.3.4 Contribution of imperfect digitization

As a last check, we estimate the contribution of the imperfect digitization on run Dec16B. This is the longest run, and the only one that allows us to decorrelate 9 time series at the first frequency, still satisfying the condition $M_f \geq p$. We find full compatibility with the previous results.

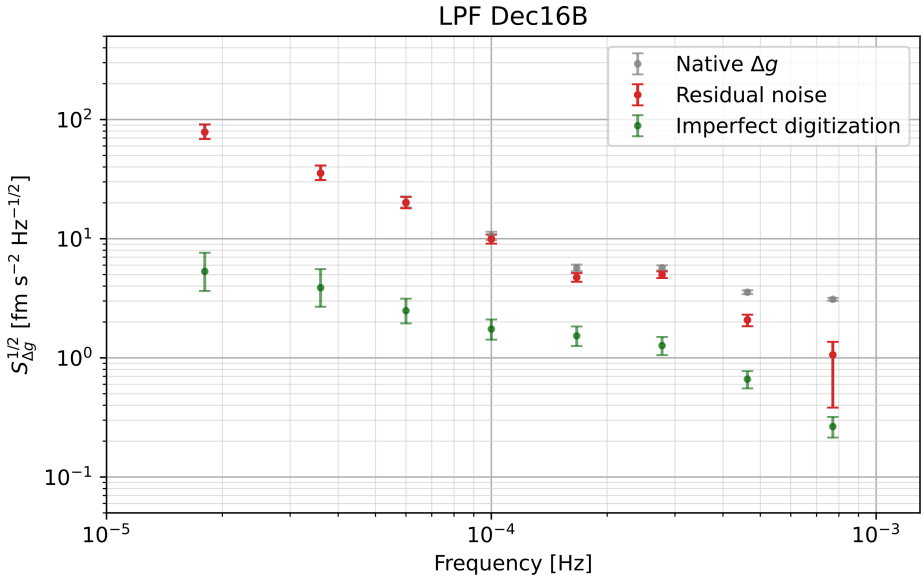


Figure 6.10: Check of the validity of the full-band result, run Dec16B. We decorrelate the subset of 9 time series of imperfect digitization. We find full compatibility with results in the complete case.

6.4 First bin decorrelation, thermal effects

We now analyze the first frequency bin, 18 µHz, separately. We noted that it is interesting for the influence of the temperature time series, hence we only focus on the decorrelation of temperature from it. There are a few reasons why it is better to deal with this bin separately. The main reason is that, as we noted, time series other than temperature have a negligible influence on Δg, especially at the lowest frequency, so we do not need to include them in the analyses. This is a key point, since for the single-frequency analysis it is necessary that the number of available periodograms exceeds the number of decorrelated time series, and the number of periodograms available at the first frequency is just a few units.
Additionally, decorrelation of a single-frequency bin can be done analytically, without needing to resort to MCMC numerical methods. Assuming a real-valued susceptibility, we use the result in Eq. (3.53).

Before decorrelating the mean temperature, we subtract the time series $-\omega_d^2 \Delta x_{\mathrm{GRS}}$ and $\kappa_t \kappa_b \Delta g_F$, which have deterministic effects on low-frequency values. However, a posteriori, we note that their influence is not as relevant as temperature. As a last note, we point out that there is no need to explicitly subtract Brownian noise, since it is negligible at this frequency.

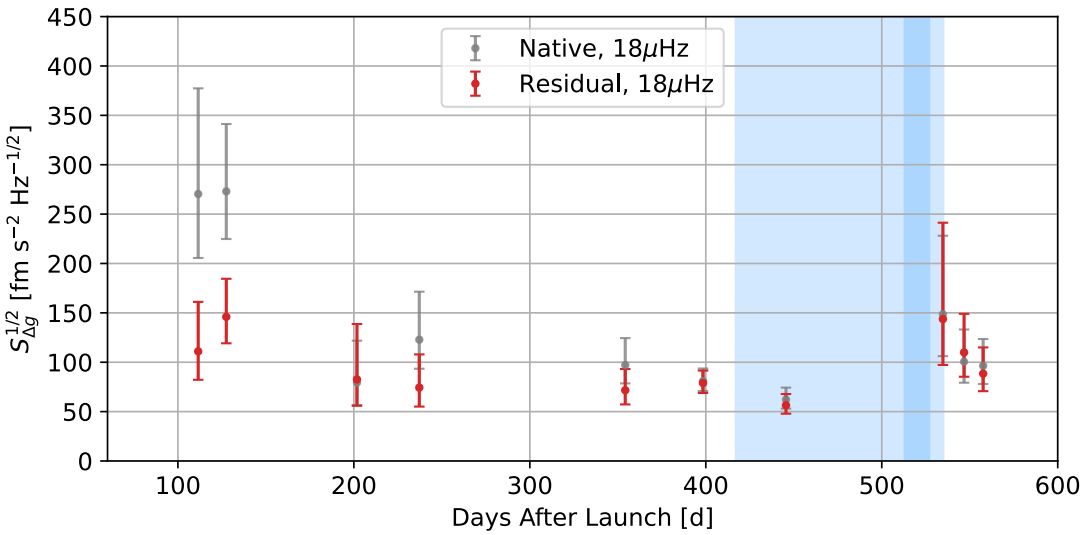


Figure 6.11: First frequency bin (18 µHz), measured native PSD, and PSD after decorrelation of the mean temperature. Only time series with more than 1 periodogram are shown (hence, longer than 3.85 d), because of numerical reasons. In runs Jul16C and May17E, the residual posterior is wider than the native $S_{\Delta g}$, because just $M \leq 4$ periodograms are available. After decorrelation, residuals are broadly compatible with a unique excess, $S^{1/2}_{\Delta g,\mathrm{decorr}}(18\,\mu\mathrm{Hz}) = 80^{+60}_{-20}\,\mathrm{fm\,s^{-2}\,Hz^{-1/2}}$. Blue spans refer to colder runs.

In Fig. 6.11, we show the ASD $S^{1/2}_{\Delta g}(18\,\mu\mathrm{Hz})$, before and after temperature decorrelation, over the course of the LPF mission. We find that the decorrelation is more effective at the beginning of the mission, a few months after venting, when pressure was the highest. After decorrelation, residuals are compatible with a unique excess, $S^{1/2}_{\Delta g,\mathrm{decorr}}(18\,\mu\mathrm{Hz}) = 80^{+60}_{-20}\,\mathrm{fm\,s^{-2}\,Hz^{-1/2}}$.

6.5 Non-synchronous time series – noise projection

The last step in the analysis of the sources contributing to Δg noise is the estimation of the noise generated by quantities that are not related to synchronously measured time series. We give estimates of the contributions from these sources, either based on other LPF experiments, or on on-ground estimates. Depending on the case, we give estimates, or worst-case estimates. Eventually, we compare the estimates with the excess noise found by the decorrelations in the previous section, and estimate the total explained noise power fraction.

First, we give a list of the effects that we consider here. In Figure 6.12, we provide our best estimate for the reference run Feb17B.

Actuation gain fluctuation noise. Among all the effects taken into account, the major source of noise in the LISA frequency band is gain fluctuation in the x-face electrode voltages. The estimate, widely discussed in [23], is represented with blue points. The estimation comes from an MCMC analysis performed on four 3-day runs in May 2016, specifically runs May16C-D-E-F, performed with different actuation configurations to isolate single-electrode contributions. Hence, the represented points are an estimate of the actuation noise level in May 2016 with UURLA authority settings; this is however well-applicable to the entire mission, as it is the best estimate available, and the actuation authority is the same in runs that we consider.
In terms of the quantity $\widetilde{S}^{1/2}_{\Delta g}$, defined in Eq. (4.4), actuation contributes with $\widetilde{S}^{1/2}_{\rm act} = (0.50 \pm 0.03)\,\mathrm{fm\,s^{-2}\,Hz^{-1/2}}$ in the frequency band [0.1 mHz, 1 mHz], accounting for $(30 \pm 4)\%$ of the measured excess noise.

Random charging noise. Random TM charging from cosmic rays is listed among the noise sources, contributing as [26]

$$S_{\Delta g}(f) = \left(\frac{\Delta_x}{MC_T}\frac{\partial C_x}{\partial x}\right)^2 S_q(f), \tag{6.19}$$

where $S_q(f) = (2e^2\lambda_{\rm eff})/(4\pi^2 f^2)$ is the shot noise due to the Poissonian charging process, presumably quite stable over the mission. What varies in time is the effective DC potential Δ_x, which depends on the compensation level in each run. In Feb17B, the compensation vas maximal, therefore the contribution of random charging (yellow line) was minimal. Even in the worse non-compensated cases, the estimated PSD of random charging is negligible with respect to the measured $S_{\Delta g}$.

Voltage fluctuation noise. Analogously to the case of random charging noise, electrode voltage fluctuations couple to the TM rms charge, as

$$S_{\Delta g}(f) = \left(\frac{q_{\rm eff}}{MC_T}\frac{\partial C_x}{\partial x}\right)^2 S_{\Delta_x}(f), \tag{6.20}$$

where $q^2_{\rm eff} = q^2_{\rm rms,1} + q^2_{\rm rms,2}$. The UV light charging system allowed to keep TM charge as low as possible, making the contribution from this effect negligible. The noise $S_{\Delta_x}(f)$ is estimated in [23], and the corresponding Δg noise is represented with orange points in Fig. 6.12.

Laser radiation pressure noise. The last effect that we take into account is the radiation pressure generated by the interferometric measurement itself. The measurement beam reflects off TM1 and TM2, with nominal power (respectively) $P_1 = 2.4\,\mathrm{mW}$ and $P_2 = 1.2\,\mathrm{mW}$, inducing outward forces on the test masses. Such forces translate into a negative differential acceleration:

$$\Delta g(t) = -\frac{2}{Mc}\left(P_1(t) + P_2(t)\right) \tag{6.21}$$

Any fluctuations in the laser beam power would then convert into acceleration noise, with an expected coupling coefficient $\sim 3\,\mathrm{pm\,s^{-2}\,mW^{-1}}$. If the power hitting the TMs fluctuated as $0.3\,\mu\mathrm{W\,Hz^{-1/2}}(1\,\mathrm{mHz}/f)$, it would account for the entire observed excess noise. However, LPF lacked a direct measurement of the optical power reflected off the test masses, which is the power of the measurement beam. Instead, as noted in reference [75], measurements combining the reference and measurement beams are the only available time series. An additional difficulty arises from the discrimination of the contribution of the two different light polarizations: in particular, the test masses are hit by a substantial fraction of parallel-polarized light, which fluctuates in power at low frequencies, and is not controlled as much in the power control loops. In [75], a worst-case estimate of the acceleration noise in run Feb17B is provided. As a dashed pink line in Fig. 6.12, we represent the worst-case contribution of laser radiation pressure, as the 95% quantile of the estimated PSD.

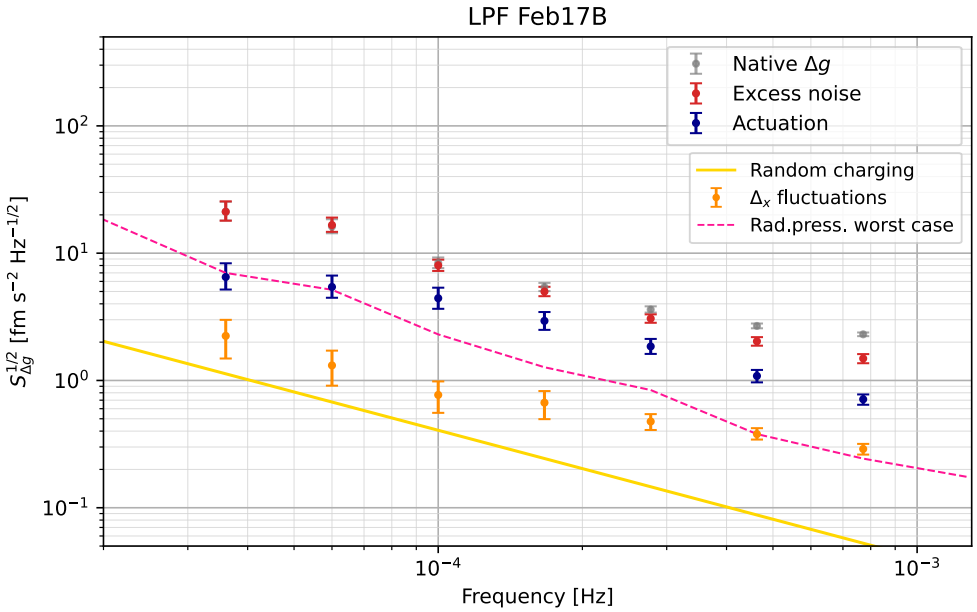


Figure 6.12: Best estimate of the contribution to differential acceleration from non-decorrelable effects. Grey points represent the measured Δg; red points are the estimate of the excess, after subtraction of the Brownian noise and the time series decorrelation in the previous section. Blue points represent the estimated contribution of actuation, as in run May16C. The orange points and the yellow line are the estimates of electric noises, respectively from TM charge variation and electrode voltage fluctuation. The pink, dashed line, is the worst-case estimate of the laser radiation pressure, adapted from [75].

6.6 Considerations on the in-band noise projection

In addition to run Feb17B, we also show the noise projection in LPF's long runs Apr16A, Nov16B, Dec16B, and Jun17 (Jun17A+Jun17B), for which we evaluated the excess noise after time series decorrelation. Projections are shown in Fig. 6.13. Orange points and yellow lines represent the estimates of the contribution from Δ_x fluctuations and random charging, based respectively on the rms TM charge and the TM DC voltages in the respective runs. Nevertheless, the estimation of the actuation gain noise (blue points) and of laser radiation pressure (dashed pink line) refer to run Feb17B.

As for run Feb17B, we note that the major known contributor to sub-mHz forces is the actuation gain noise, contributing up to 40% power at about 0.3 mHz, and less than that at the other frequencies. Temperature affected just the first bin, but did not have any influence on the following ones. Electrical effects (Δ_x fluctuations and random charging), have negligible impacts thanks to the employed compensation strategies, which ensured low contributions from these effects. The contribution of laser pressure might in principle be relevant, but worst-case estimates, at least in the reference run Feb17B, show that a contribution higher than 10% power is implausible. Nevertheless, none of the sources taken into account can clearly explain the detected PSD below 0.1 mHz, nor its f^{-2} dependence.

Additional discussions on possible sources, in a certain sense more qualitative, are provided in the discussion Chapter 8.

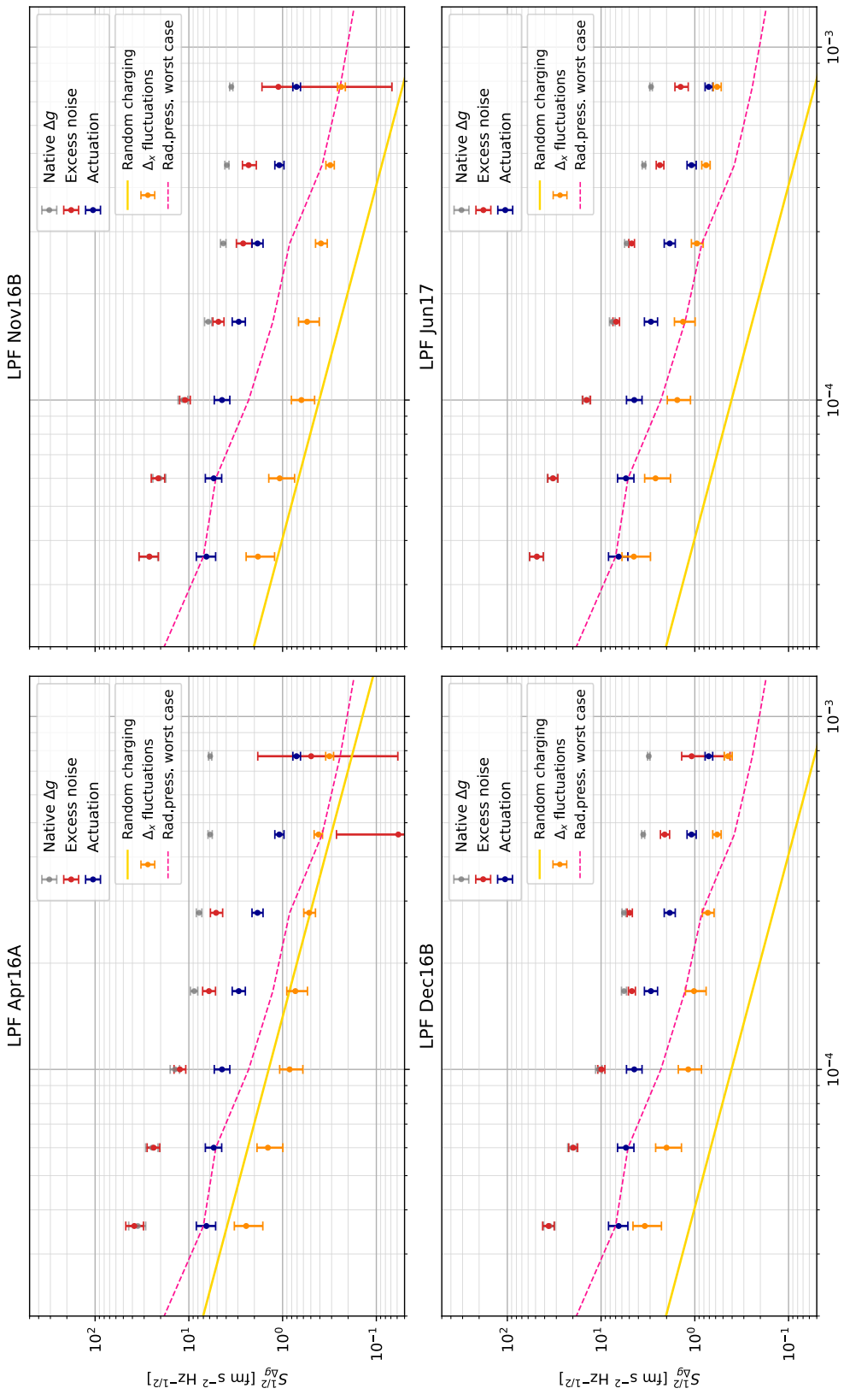


Figure 6.13: Best estimate of the contribution to differential acceleration from non-decorrelable effects, long LPF runs Apr16A, Nov16B, Dec16B, Jun17B. Legend is the same as Fig. 6.12. The estimates of actuation noise (blue points) and radiation pressure noise (dashed pink line) refer to run Feb17B.

7 LPF glitches

Phenomenology, properties and source hypotheses

This chapter is dedicated to the spurious transient events detected during the LPF operations, during noise-only science runs. These events, commonly known as *glitches*, spoilt the measurements of the quasistationary noise, resulting in an effective overestimation of the PSD if not subtracted. These glitches[1] were detected in all of the steps of the post-processing chain presented in Sec. 2.2, from raw interferometer data to the out-of-loop differential acceleration $\Delta g(t)$, even after cleansing from inertial forces. An example of glitches detected during the Dec16B noise run is depicted in Fig. 7.1. Some more examples are given later.

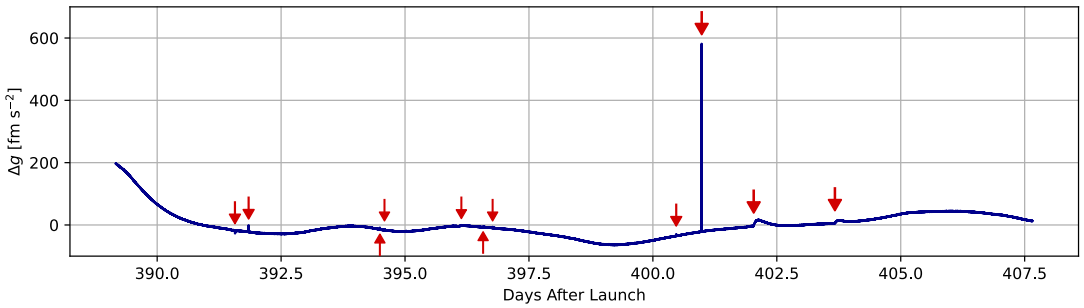


Figure 7.1: Example of glitches during the Dec16B noise-only run. During about 18 days, 11 glitches were detected. All of these glitches belong to the impulse-carrying family, as described later in text. Data have been low-pass filtered with a 100 s Blackman-Harris windowing function. We notice that the glitch occurring at DAL 401 is the most powerful glitch detected, peaking at (8.78 ± 0.01) pN force.

In this chapter, we first show that glitches can effectively be divided into two different categories (Sec. 7.1), with different properties and different natures, and give a physical explanation for one of the two categories. We thoroughly analyze the detected properties and parameters (Sec. 7.2), the time-of-arrival statistics, and the measured durations and amplitudes. We attempt to understand the real nature of such events and the source causing them: first, we cross-check dynamical variables (Sec. 7.3), stating the different nature of the two families. Then, in Sec. 7.4, we look for coincident events in environmental parameters (temperature, magnetic fields, etc.). We rule out many possible explanations, and discuss the implications of our analysis toward the understanding of the physical origin of glitches (Sec. 8.3).

Regarding the perspectives for LISA and further analyses, we investigate the

[1] In particular, the most powerful ones.

possibility of detecting such glitches in ground-based experiments, especially at the torsion pendulum facilities located at the University of Trento (Sec. 7.5).

The results presented in this chapter, and in the glitch section of the next one, Sec. 8.3, have been published in [4]. In that work, we completed and extended the results presented in [1–3]. Most of the figures in this chapter and the next are adapted from there, and some passages may be similar or verbatim.

7.1 LPF glitch classification

Before going into the details of the glitch analysis, some remarks are needed about the data runs that we have used for this analysis.

1. We have used all data runs listed in Table A.1, comprising both the noise-only ones and some other runs during which the satellite was in steady control conditions. We comprise a run with a low-amplitude injected sinusoidal signal (100 fN), which does not affect our analyses.

2. We also included some short-lasting runs, in addition to the more significant week-long ones, for statistical purposes.

3. We note that we included both nominal-temperature, low-temperature, and cold-temperature runs. As specified in Sec. 2.1.6, we recall that we refer to *ordinary runs* (OR) as the ones performed at 22 °C or 11 °C. We refer to *cold runs* (CR) as the ones performed in May 2017, at about 0 °C. This distinction is crucial in these analyses, since the unstable state detected in CR led the glitch rate to a high increase. The majority of glitches have been detected in CR.

4. We also include a run with grabbed TM, to gain insight into the nature of glitches.

Let's thus begin with the description of the glitch templates and the glitch subdivision into two separate families. We analyze glitches as measured in the $\Delta g(t)$ time series. We lowpass the data for the analysis. For filtering, we use a 100-s Blackman-Harris window, convoluting it over the data series. The window length is chosen as required by different cases, since it directly affects the roll-off frequency of the filter. This provides us with a reliable tool for assessing the sign properties of glitches: this kind of filter is indeed sign-preserving, and does not introduce any alterations of the signal sign, nor oscillations; moreover, it preserves the total integral.

We recall that $\Delta g(t)$ is the differential out-of-loop acceleration between the test masses. Given the system dynamics, assuming that glitches are caused by forces, there is no way of discriminating which of the two masses is acted on. The only property that can be deduced is if it acted inwards ($\Delta g > 0$) or outwards ($\Delta g < 0$).

In Figure 7.2, we show two examples from two different glitch families, with completely different time profiles and durations, which we call *impulse-carrying glitches* (ICG) and *fast, low-impulse glitches* (FLG). In the following, we make a complete distinction between the two cases. We identify a few exceptions that do not fit into these categories, that we treat separately.

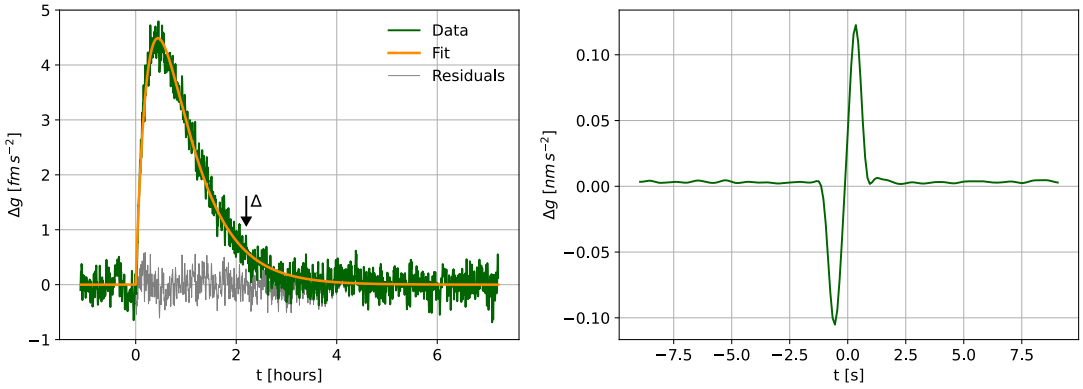


Figure 7.2: (left) Example of an impulse-carrying glitch. The picture shows: (green) the native data after low-pass filtering and background subtraction; (orange) the fitting template; (gray) the residual after subtraction of the template. The represented event corresponds to an impulse per unit mass $\Delta v = (21 \pm 2)\,\mathrm{pm\,s^{-1}}$, and a duration $\Delta = (7500 \pm 200)\,\mathrm{s}$, as indicated by the black arrow and defined later in text.
(right) Example of a fast, low-impulse glitch. The figure shows data after the 0.5 Hz low-pass filtering. This event could be seen even without filtering data, as a series of 2-5 oscillating experimental points (see text).

7.1.1 Impulse-carrying glitches

We detected 432 events belonging to the ICG category (98 in ordinary runs, and 334 in cold runs). The analysis was made manually, following what was already done in [3]. Extending what was found in [3], we find that glitches within this category are shaped as in the following two templates:

$$h_1(t) = \frac{\Delta v}{\tau^2}\, t'\, e^{-t'/\tau}\, \Theta(t'), \quad t' = t - t_0 \tag{7.1}$$

$$h_2(t) = \frac{\Delta v}{\tau_1 - \tau_2} \left(e^{-t'/\tau_1} - e^{-t'/\tau_2} \right) \Theta(t'), \quad t' = t - t_0 \tag{7.2}$$

Where t_0 is the occurrence time, $\Theta(t)$ is the Heaviside step function, Δv is the total transferred impulse per unit mass, and the τ parameters rule the duration of the events. Precisely, Eq. (7.2) had already been used in [2, 3]: the larger between τ_1 and τ_2 sets the decay time, while the shorter sets the rise time; Eq. (7.1) is the limit of Eq. (7.2) in case $\tau_2 \to \tau_1 = \tau$, and effectively removes the need of the delta function $\Delta v\, \delta(t')$ as in [3], with the advantage of having a time structure[2], and fitting data effectively. Moreover, [3] noted a high anti-correlation in many cases between τ_1 and τ_2, resulting from the fit procedure. It hints indeed a certain degree of overfitting, and also suggests that modeling the system with two different τs does not add any information to a single-tau model.

The main property of these templates is that the acceleration profile is either positive or negative, and leaves no residual acceleration after its occurrence. The template

[2] Additionally, note that this template corresponds to the first-order shapelet considered in [76]

integral, Δv, represents the total transferred impulse per unit mass[3], $\int_0^\infty \Delta g(t)dt = \Delta v$. From a frequency-domain perspective, Δv is also the zero-frequency limit of the Fourier transform of the $h_1(t)$ and $h_2(t)$ templates. Noticing this is relevant, since it implies that ICGs have a strong low-frequency component; the Fourier transforms of the templates are indeed

$$h_1(\omega) = \frac{\Delta v e^{-i\omega t_0}}{(1+i\omega\tau)^2} \quad h_2(\omega) = \frac{\Delta v e^{-i\omega t_0}}{(1+i\omega\tau_1)(1+i\omega\tau_2)} \tag{7.3}$$

Where $\omega = 2\pi f$. We notice that, in the h_1 case, $|h_1(\omega)| \propto \omega^{-2}$ at frequencies $\omega \gtrsim 1/\tau$.

Glitches of this family could be fitted well with the template above, as represented in Fig. 7.2 on the left, leaving a residual that is by-eye comparable to the absence of the glitch. To quantify this statement we compare, in Fig. 7.3, the ASD of the glitch-free stretches (green) to the ASD of the entire Feb17B run after glitch removal. We first note that the glitch subtraction leaves no residual in the stretch PSD; moreover, we note that it allows to analyze lower frequencies, whose analysis is otherwise prevented by the shortness of the stretches.

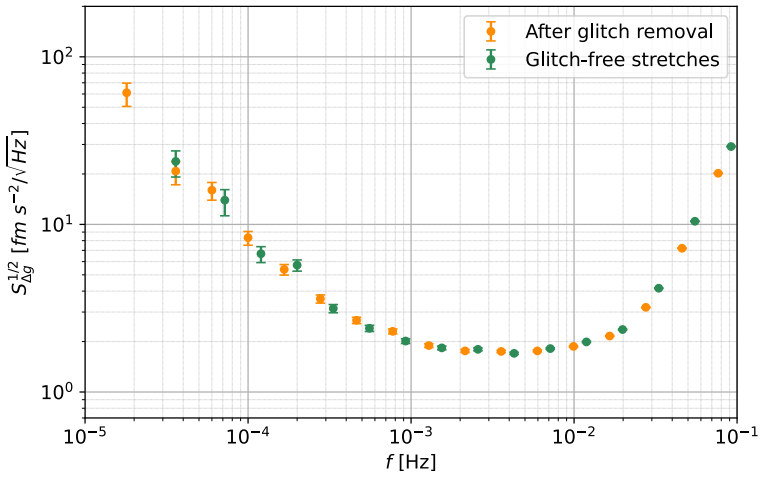


Figure 7.3: Amplitude spectral density (ASD) $S_{\Delta g}^{1/2}(f)$ of quasistationary noise, vs the frequency f, of the Feb17B noise run. Green data points: ASD of glitch-free stretches only. Orange data points: ASD of residuals after glitch removal. Note that frequencies are slightly shifted, because of the frequency choice algorithm.

The subtraction workflow is the following:

1. We filter data convoluting with a 100 s-long Blackman-Harris window, which has an effective roll-off frequency at 10 mHz. The filter suppresses the intense high-frequency noise coming from the double time derivative of the interferometer readout noise, and allows to effectively identify by eye the presence of glitches.

2. Subsequently, for each glitch, we fit a polynomial to the data in two 1000 s-long stretches, one immediately preceding the glitch stretch, and the other

[3]This definition assumes that the glitch is a real force on the TM(s), implicitly ruling out the case that the glitch is an interferometer readout error. We will prove (Sec. 7.3.1) that this is indeed the case.

immediately following it. We choose a first-order polynomial for glitch stretches shorter than 1000 s (see later for the definition of duration), and second-order for longer ones. We then subtract the best-fit polynomial from data. The result of this procedure, for one of the longest glitches, is shown in Fig. 7.2, left panel. Such background subtraction was necessary to get rid of the long-term drift that affected all data, mostly due to the gravitational signal from propellant tank depletion and long-term temperature variations. The use of a second-order polynomial for the longest glitches was able to accommodate some drift rate variation over many hours duration.

3. We low-pass the templates in Eqs. 7.1 and 7.2 with the same filter we used for data, and attempt fitting with the two templates. Among the two fits, we choose the one which leaves less residual power.

With this approach, we find that, of the 432 analyzed events, 48 can be fitted with the two-tau model (Eq. (7.2)), and 384 with the single-tau one (Eq. (7.1)). The statistics concerning these parameters are analyzed later.

To properly define a duration parameter, agnostic of the fitting template, we calculate the effective duration Δ, defined as the time span, from the beginning of the glitch, comprising 99% of the total signal energy,

$$\Delta : \frac{\int_0^{\Delta} h^2(t)dt}{\int_0^{\infty} h^2(t)dt} = 99\%. \tag{7.4}$$

The duration Δ is also reported in the example in Fig. 7.2. In the simple single-tau case, Eq. (7.1), it corresponds to $\Delta \sim 4.20\tau$. Errors on Δ are found as explained in Appendix C.

No glitches of this kind have been found during the run with grabbed TM; however, it was too short to draw any conclusions.

7.1.2 Fast, low-impulse glitches

The second population of glitches is characterized by short duration and minimal total impulse. Very few examples of these events are found in ordinary runs: only 4 during the analyzed ~ 138 days. Glitches of this kind occurred numerously during the cold runs, amounting to 152 events[4]. In Fig. 7.4, we provide as an example the data from the first cold run, May17A.

There are many relevant differences with glitches belonging to the first family:

1. They carry no significant impulse (per unit mass) $\Delta v_{\rm glitch} = \int_0^{\infty} \Delta g(t) \mathrm{d}t$, hence they have no detectable counterpart in the feedback force time series $g_{\rm c}(t)$. The frequency spectrum of these events has a strong high-frequency content, so no lowpass filter is usually required to detect them. Furthermore, the decorrelation of inertial forces induces a high-frequency noise, so FLGs are even better visible before that correction.

[4]Detection was basically made manually, with the help of a few computational tools. It is not to exclude that the true number might be slightly different from 152, but this does not make a relevant difference in our discussion.

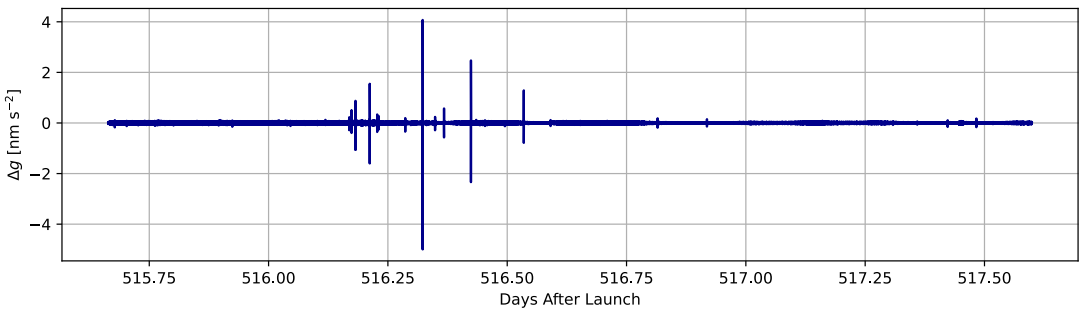


Figure 7.4: Example of glitches during the May17A noise-only cold run, unfiltered. Only FLGs are visible by eye, a filter and a scale reduction would be necessary to identify the ICGs.

2. The overall duration of the glitch is compatible, within errors, with that of the convolution of the impulse response of the filter used to estimate the second time derivative, with that of the low-pass filter. Without applying any filters, FLGs show the characteristic 5-point shape of the numerical derivative.

3. Despite what was noted in the first point, we apply a low-pass filter with a roll-off frequency of $\simeq 0.5\,\mathrm{Hz}$ (Blackman-Harris, 2 s long), which is needed to detect some events in run May17B.

4. Some glitches of this kind have also been detected during the run with grabbed TMs.

In a first search for glitches, 2 of this kind were found in ordinary runs and subtracted from data by fitting to the second derivative of a step function, properly filtered. However this procedure had no visible effect on the PSD estimation, as this glitch category only impacts high frequencies. Therefore, the two additional glitches of this kind that we found in a second search, were not subtracted. We did not subtract this kind of glitches in cold runs, as the fitting turned out to be unfeasible in most of the cases. Thus, in the following analyses, we will consider just two parameters defining FLG glitches: the impulse $\Delta v_{\mathrm{glitch}}$, and the time of occurrence defined as the time when $\Delta g(t)$ reaches the maximum absolute value.

7.1.3 Exceptions

We found 5 events, belonging to the ICG family, which did not resemble the templates in Eq. (7.1) and (7.2). Two of these glitches, the longest ones and both in ordinary runs, required a third exponential function to account for some fine structure of the glitch onset, and to allow for good-quality glitch subtraction. The fit could however be performed with the template in Eq. (7.2), but adding a third exponential function allowed a better resemblance to the model:

$$h_3(t) = \Delta v \left(\frac{\tau_1\, e^{-t'/\tau_1}}{(\tau_1-\tau_2)(\tau_1-\tau_3)} + \frac{\tau_2\, e^{-t'/\tau_2}}{(\tau_2-\tau_3)(\tau_2-\tau_1)} + \frac{\tau_3\, e^{-t'/\tau_3}}{(\tau_3-\tau_1)(\tau_3-\tau_2)} \right) \Theta(t') \tag{7.5}$$

We also found 3 glitches that still carry a significant impulse, but, contrary to the others, do cross $h(t) = 0$ during their evolution, even though only once. These

glitches are fitted to

$$h_c(t) = h_2(t) + \tau_d \dot{h}_2(t), \tag{7.6}$$

with τ_d is a constant that can have any sign. An example of these events is shown in Fig. 7.5.

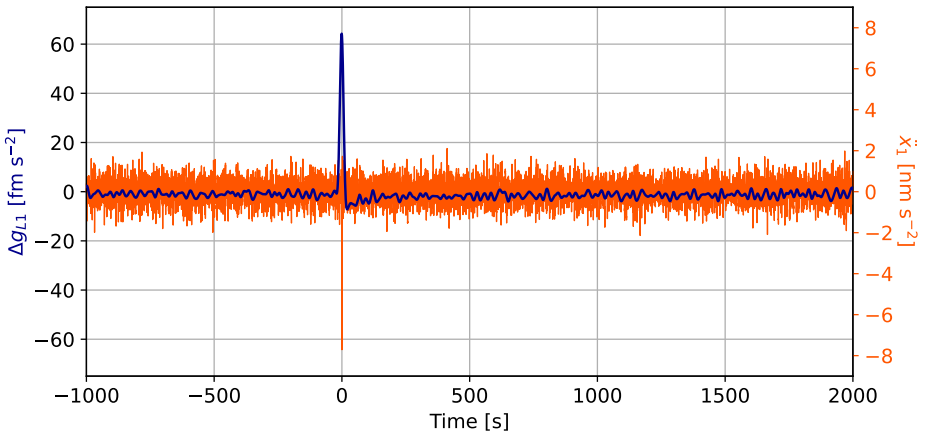


Figure 7.5: Lowpassed (50 s Blackman-Harris windowing) anomalous glitch occurring in the Jun17B run, responding to the zero-crossing template in Eq. (7.6), On the left scale, Δg_{L1} is Δg before the correction of the tangential Euler force. We also show the second time derivative of the raw readout $x_{1,\mathrm{OMS}}$, which will be used later in the discussion (orange, right scale).

7.1.4 Other spurious signals in the data

We also found a few events in $\Delta g(t)$ caused by the impact of the spacecraft with micrometeoroids, which is already a well-known phenomenon for LPF [41]. Even though $\Delta g(t)$ has been corrected for the acceleration of the spacecraft, some very energetic events might leave a residual signal in the data, likely because of unperfect calibration. Glitches of this kind are however well understood.

We also found some spurious events, spikes, or noisier data stretches, observed when operating particular devices, that could be suppressed by turning their source off. We do not discuss any further events of this kind.

7.2 Glitch parameter statistics

We now proceed with the analysis of the glitch parameter observed statistics, their time of arrival, amplitude, and duration. This section mainly focuses on the ICGs, due to their prominent impact on mission performance, the possibility to fit them, and their higher number of parameters. At the end of this section, we also treat the FLGs and the few exceptions with the template in Eq. (7.6).

7.2.1 Occurrence rate, impulse-carrying glitches

Let's start with the occurrence glitch rate, starting with ICGs. We define the waiting time ΔT as either the time between two subsequent glitches, or the time between the

starting moment of the run and the first glitch, relevant for very short runs with few glitches. We show the waiting time histograms, distinguishing ordinary runs and cold runs, in Fig. 7.6.

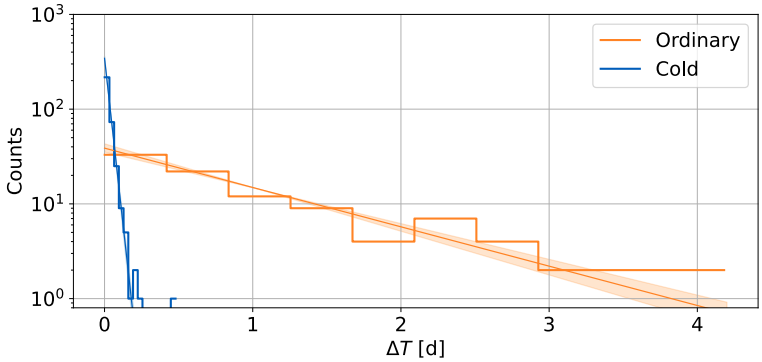


Figure 7.6: Histogram of the waiting time ΔT for ordinary runs (orange) and cold runs (blue). Straight lines, and associated shadowed areas, represent, respectively, the Bayesian fit to an exponential distribution, and the corresponding 1σ confidence interval. For the cold runs, the fit is limited to data with $\Delta T \leq 0.2\,\mathrm{d}$. The rates from the Bayesian estimation for ordinary runs, and cold runs with $\Delta T \leq 0.2\,\mathrm{d}$, are, respectively, $\lambda = 0.96^{+0.11}_{-0.09}\,\mathrm{d}^{-1}$ and $\lambda = 32^{+2}_{-2}\,\mathrm{d}^{-1}$.

We test the hypothesis that the occurrence times follow a Poisson distribution, testing that the waiting times follow an exponential distribution,

$$p(\Delta T|\lambda) = \lambda e^{-\lambda \Delta T} \tag{7.7}$$

We employ the Lilliefors test [77], which tests data against the null hypothesis of data being exponentially distributed. For both ordinary and cold runs, the result is compatible with ΔT being exponentially distributed. This indicates that the occurrence distribution is Poissonian, and hence the events happen independently. Data for ordinary runs are well fitted to an exponential distribution with an average rate $\lambda = 0.96^{+0.11}_{-0.09}\,\mathrm{d}^{-1}$ ([5]).

Data for the cold runs, and for $\Delta T \leq 0.2\,\mathrm{d}$, are also compatible with an exponential distribution with a much higher rate $\lambda = 32^{+2}_{-2}\,\mathrm{d}^{-1}$. The distribution shows some excess counts at longer times, see Fig. 7.6, which originate from temperature variations at the end of the cold runs (see below).

Bayesian waiting time estimation

In Fig. 7.6, we presented a joint analysis of the glitch arrival times, implicitly assuming that glitches occurred at a constant rate during the entire mission. This is a hard statement, which needs further testing, and stability analysis of the glitch rate. To address it, we first split the time series into 12 groups, grouping runs whose starting

[5]This rate of $\lambda = 0.96^{+0.11}_{-0.09}\,\mathrm{d}^{-1}$ is apparently slightly higher than that reported in our preliminary search [2] of $\lambda = (0.78 \pm 0.02)\,\mathrm{d}^{-1}$. We have traced back this apparent discrepancy to the smaller subset of runs used in [2], and to a mistake in reporting the error. The event rate estimated with the current Bayesian analysis, at 1σ confidence level on the same subset of runs, gives $\lambda = 0.75^{+0.13}_{-0.09}\,\mathrm{d}^{-1}$, which is compatible with the current estimation at 1σ confidence level.

times differ by less than a month (see Fig. 7.7), making subsets of arrival times. Each subset contains about 10 events.

Within each subset, we run a Lilliefors test to make sure that data are compatible with an exponential distribution, and infer a distribution of the theoretical event rate λ. The test is always successful, though we note that the number of events in each subset is really low. Under the assumption that each event is an independent extraction from an exponential distribution with rate λ, we apply Bayes' theorem to the joint probability of n events to get its posterior distribution, with a uniform prior on λ:

$$\mathcal{L}(\lambda) = \frac{S^{n+1}}{\Gamma(n+1)}\lambda^n e^{-S\lambda}, \quad S = \sum_i \Delta T_i \tag{7.8}$$

We estimate λ as the probability-maximizing value, and also provide the asymmetric error bounds at 1σ confidence level. The result, shown in Fig. 7.7, is compatible with the fit of the waiting time distributions in Fig. 7.6.

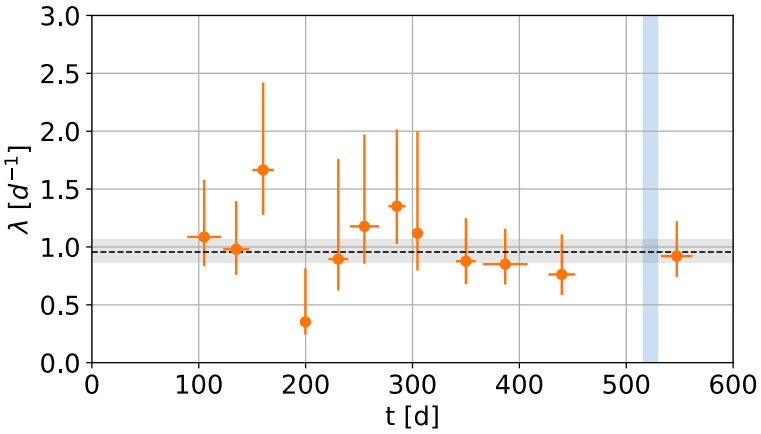


Figure 7.7: Glitch occurrence rate λ, during ordinary runs, vs time t from launch, over the LPF mission. Points are calculated by grouping glitches observed during runs with starting times differing by less than a month. Vertical and horizontal error bars are, respectively, our Bayesian estimates at 1σ confidence level, and the time span of the considered epoch. The dashed line and the gray shaded area represent the mean rate and its error from the Bayesian estimate. The blue shaded area indicates the epoch of cold runs, which we analyzed separately.

Cold runs performed in May 2017 showed a completely different situation, with a reversible increase in the ICG glitch rate. The temperature cooldown, from 11 °C to the target 0 °C. began on April 29th and was completed on May 2nd. Unfortunately, the target temperature was below the lower measurement threshold of the EH thermistors, thus we need to use instead the thermistors located on the LPT bay.

On May 2nd, two days after the beginning of the cooldown, the noise-only run May17A started; temperature was however still slightly decreasing, as can be seen in Fig. 7.8, red points and right scale. Temperature was kept constant for about 10 days, and, during run May17D, it was increased back, targeting 11 °C. As can be seen in the figure, temperature had risen by about 5 °C, when the measurement stopped.

As for the ordinary runs, we split the cold runs into subsets, $\sim$ 23 h-long, and analyze the rate evolution in such subsets. The results are shown in Fig. 7.8, blue

points and left scale. Note that the rate variations at the end of run May17D fully explain the few excess counts at longer waiting times in the cold runs histogram of Fig. 7.6.

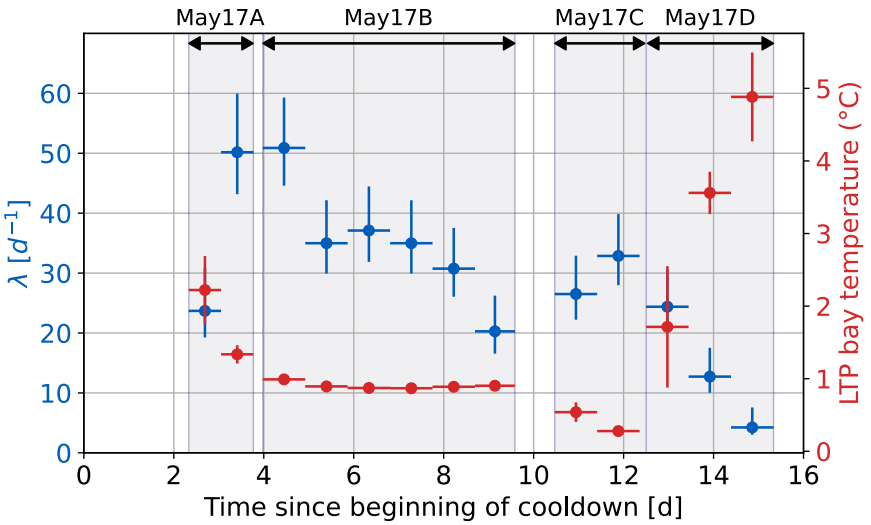


Figure 7.8: Glitch occurrence rate λ (left scale, blue), and LTP bay temperature (right scale, red), during cold runs, as a function of time since the beginning of cooldown. Small temperature changes before day 12 were due to adjustments to the heaters' settings. Reheating to ordinary conditions started on day 12. We also recall that run May17B was abruptly interrupted, because of a sudden uncontrollable acceleration, likely of mechanical origin [45].

7.2.2 Occurrence rate, fast low-impulse glitches

Glitches in this category were too rare in ordinary runs to analyze their statistics, so we just analyze occurrences in cold runs. Figure 7.9 shows the histogram of the waiting time ΔT.

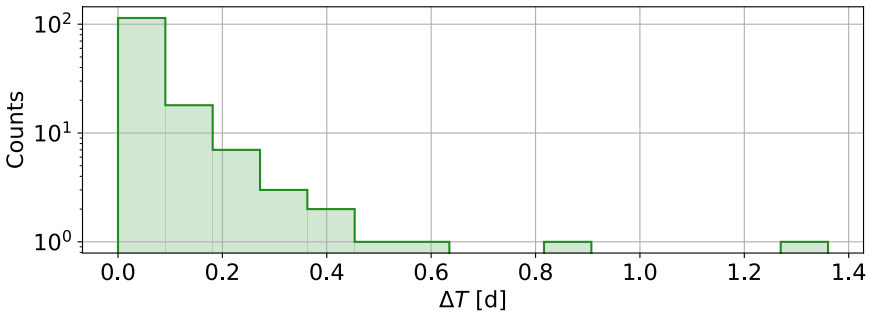


Figure 7.9: Histogram of the waiting time ΔT for fast, low impulse glitches in cold runs. The Lilliefors test proves that data do not follow an exponential distribution, hence indicating that the underlying distribution is not Poissonian, and the events are not independent.

We run a Lilliefors test to test if the arrival times followed an exponential distribution, but it failed indicating some departure from Poisson statistics. Some subsets of the glitches are indeed clustered, especially when heaters were turned on again during

May17D, as shown in Fig. 7.10. On DAL 528 during run May17D, at 11:50:00 UTC, heaters were turned on. After that, some FLG events occurred at an increased rate.

However, tests run on subsets of the data show that some subsets could still be compatible with an exponential distribution.

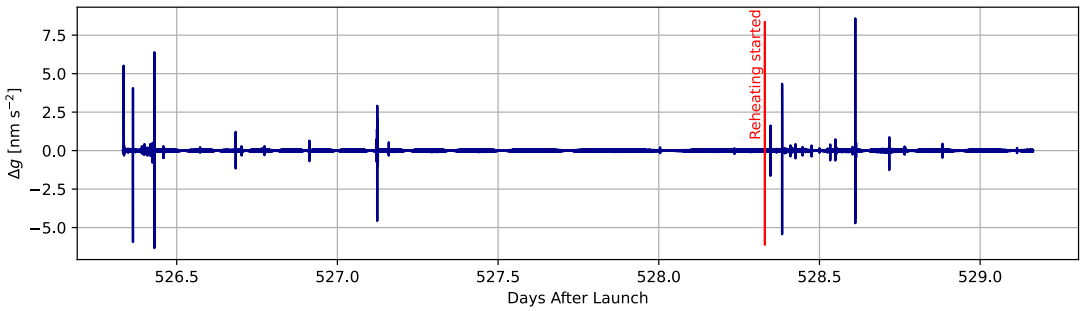


Figure 7.10: Example of glitches during the May17D noise-only cold run, unfiltered. The red line marks the exact time of reheating start, at 11:50:00 UTC on DAL 528. After that moment, it is possible to see that FLGs occurred with an enhanced non-Poissonian rate.

7.2.3 Impulse and duration, impulse-carrying

Regarding the ICGs, which were fitted to the templates in Eqs. (7.1, 7.2, 7.5), we also analyze the distributions of the glitch parameters, namely, the impulse per unit mass Δv, and the duration Δ, which condenses the information given by τ_1, τ_2, τ_3. In Figure 7.11, we represent the distributions of such parameters, distinguishing positive/negative events, and ordinary/cold runs. It is remarkable that the distributions of positive and negative events in ordinary and cold runs are quite different. We will discuss it in Sec. 8.3.

Error bars represented in the picture are not the ones that were determined via the fit, since the least-squares method does not lead to accurate estimates in presence of high, non-white noise. Rather, we estimate the errors using a frequency-domain method, in which we directly employ the measured PSD in glitch-free stretches. We prefer to explain it in detail in the Appendix, App. C. That method also allows to calculate the signal-to-noise ratio (SNR) of each glitch, with respect to the noise PSD within its run. Especially at high Δ in ordinary runs, it can be noticed a gap between the SNR $= 3$ curve and the detected glitches. We note that the line is calculated using the noise PSD of the Feb17B noise run, which is slightly different from the noise PSD in other runs. A better grasp of the SNR of glitches with different duration can be got looking at the right panel of Fig. 7.12, where we show SNR(t).

To better show the SNRs, in Fig. 7.12 on the left panel we show the distribution of the SNRs, in either ordinary or cold runs. The lowest detected SNR is ~ 3, whilst the highest is close to 10^3, associated with high-impulse short-lasting events.

We also analyze the evolution of the Δ and Δv parameters over the LPF mission. In Fig. 7.14 we represent, on the right panel, the absolute impulse per unit mass $|\Delta v|$ (top), and the duration Δ (bottom) of ICGs, as a function of time from launch. We note no apparent dependence of parameters on the time from launch. The apparent clustering is simply due to the different measurement runs. On the left panel, we

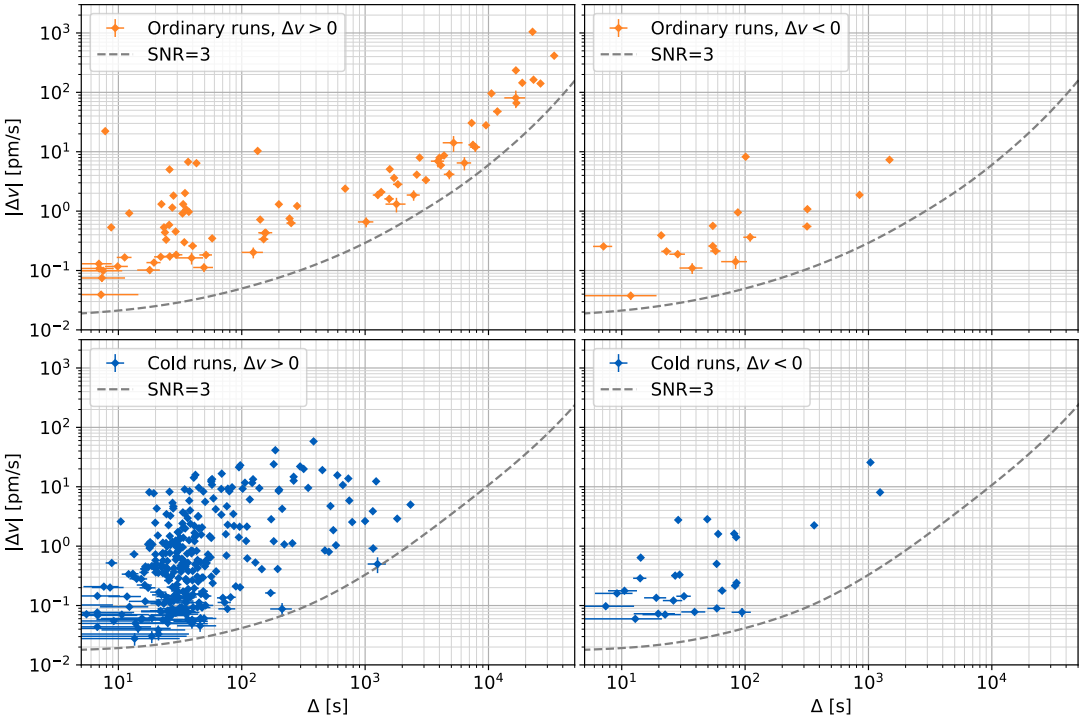


Figure 7.11: Absolute impulse per unit mass $|\Delta v|$, and duration Δ, of impulse carrying glitches fitted to templates in Eqs. (7.1, 7.2, 7.5).
Upper left panel: the 81 positive impulse glitches observed during ordinary runs.
Upper right panel: the 17 negative impulse glitches observed during ordinary runs.
Lower left panel: the 306 positive impulse glitches observed during cold runs.
Lower right panel: the 28 negative impulse glitches observed during cold runs.
The gray dashed line represents the amplitude of a glitch of the kind in Eq. (7.1) that would have SNR = 3. In the two upper panels, SNR is calculated for the lowest noise ordinary run of February 2017. In the lower panels, the line is calculated for the sensitivity of the cold runs. For a detailed analysis of the glitch SNR, see Fig. 7.12. Errors are estimated as described in App. C.

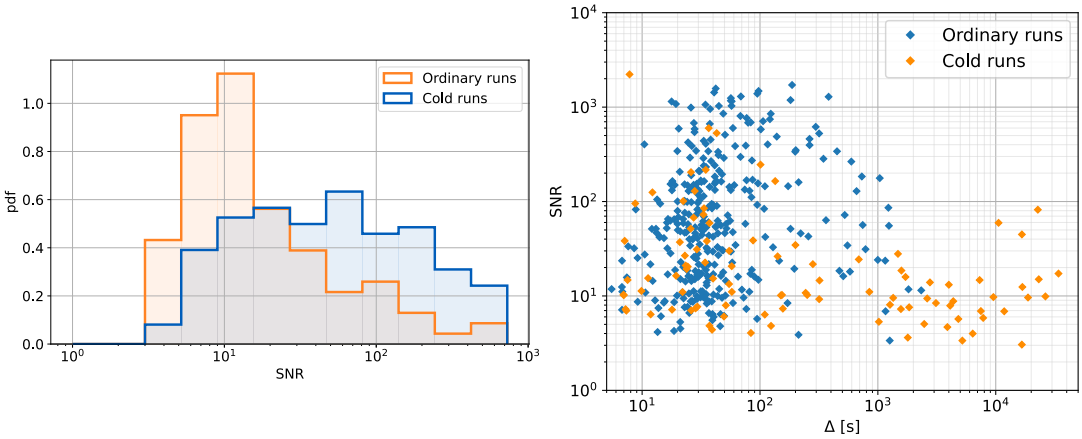


Figure 7.12: (left panel) Histogram of the SNR for all glitches of Fig. 7.11. The lowest observed value is SNR ~ 3. The probability density refers to the logarithm of SNR.
(right panel) SNR of all glitches of Fig. 7.11, as a function of their duration Δ.

represent the distribution of the measured peak $|\Delta g_{\max}|$, calculated with the templates in Eqs. (7.1, 7.2, 7.5). The very same data, divided per run and sign, are represented in Fig. 7.14.

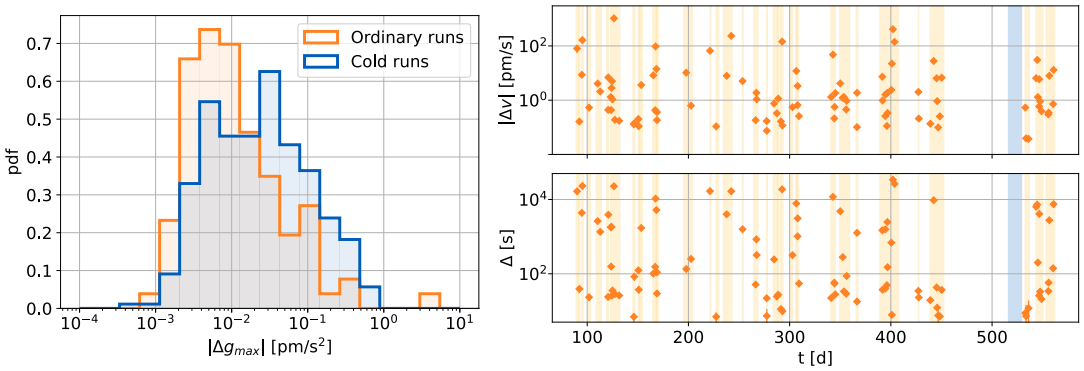


Figure 7.13: (left) Distribution of the maximum acceleration $|\Delta g_{\max}|$, associated to ICG glitches, ordinary runs and cold runs.
(right) Absolute impulse per unit mass $|\Delta v|$ (top), and duration Δ (bottom) as a function of time from launch for the ICG glitches of Fig. 7.11.

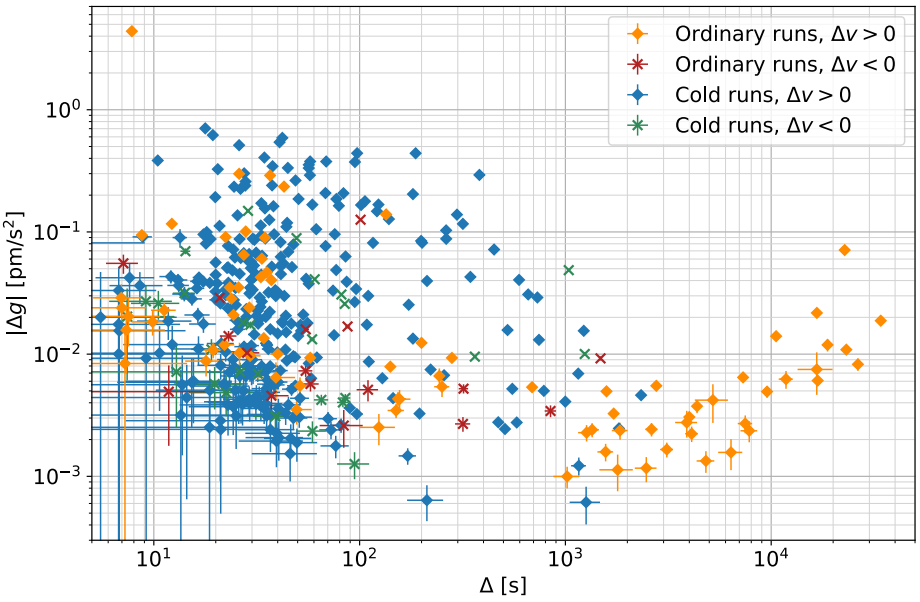


Figure 7.14: Absolute peak acceleration $|\Delta g_{\max}|$, vs duration Δ, of impulse-carrying glitches fitted to templates in Eqs. (7.1, 7.2, 7.5).

7.2.4 Impulse, fast low-impulse glitches

In Figure 7.15, we show a histogram of the impulse Δv_{glitch} of FLGs, estimated by numerically integrating the data stretches containing the glitch, and normalized to the estimated error Δv_{rms}. The latter, namely the error on the impulse estimation, is estimated, by performing the same numerical integration procedure on random segments, glitch-free, of the same length of the glitch stretch, and within the same

run. If Δv_{glitch} was normally distributed with zero mean and standard deviation equal to Δv_{rms}, then $\Delta v^2_{\text{glitch}}/\Delta v^2_{\text{rms}}$ would follow a chi-square distribution with one degree of freedom. This is the result that we find, indeed the histogram is compatible ($p = 0.23$) with the expected distribution, proving that FLGs have no discernible impulse Δv_{glitch}.

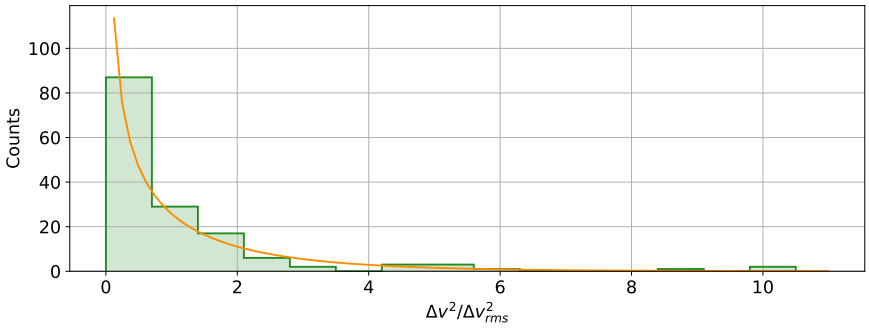


Figure 7.15: Histogram of $\left(\Delta v^2_{\text{glitch}}/\Delta v^2_{\text{rms}}\right)$ for low impulse glitches. The orange line is the normalized chi-squared distribution with one degree of freedom.

Exceptions, zero-crossing ICG glitches

The preceding figures do not include the 3 glitches that are fitted to the template in Eq. (7.6). We summarize their properties in Table 7.1.

Table 7.1: Observed glitches corresponding to the template in Eq. (7.6). Values above the horizontal line refer to ordinary runs. Values below that same line, to cold runs.

Δv [pm/s]	Δ [s]	τ_d [s]
(0.10 ± 0.01)	(35 ± 8)	(-37 ± 9)
(0.81 ± 0.03)[a]	(111 ± 3)	(72 ± 4)
(0.42 ± 0.02)	(143 ± 50)	(-29 ± 22)

[a] Corresponds to an event in $\ddot{x}_{1,\text{OMS}}(t)$, see Sec. 7.4 and Fig. 7.5 on page 139.

7.3 ICG joint analysis with dynamical time series

We analyze the coincidence between ICGs and some LPF dynamical variables, to better understand the glitch nature.

The first variable that we study is the differential TM displacement, as it is measured by both the OMS sensing and the GRS capacitive sensors. Analyzing this parameter allows to put constraints on the glitch nature, and understand if they are to ascribe to true forces or could as well be interferometer transients.

7.3.1 Joint analysis with capacitive GRS Δx

The motion of TMs along the x direction, and relative to the spacecraft, has been measured not only by the interferometer system, but also by the capacitive sensor of the GRS, as detailed in Sec. 2.2.6. We recall that, from the GRS coordinates $x_{1,\mathrm{GRS}}(t)$ and $x_{2,\mathrm{GRS}}(t)$, we formed a measurement of their relative displacement, independent of $\Delta x_{12,\mathrm{OMS}}(t)$:

$$\Delta x_{12,\mathrm{GRS}}(t) = x_{2,\mathrm{GRS}}(t) - x_{1,\mathrm{GRS}}(t) \tag{7.9}$$

Assuming perfect calibration, the difference between the GRS and OMS measurements only contains the difference between the noise terms:

$$\Delta x_{12,\mathrm{OMS}}(t) - \Delta x_{12,\mathrm{GRS}}(t) = n_{\mathrm{OMS}}(t) - n_{\mathrm{GRS}}(t). \tag{7.10}$$

Hence, this difference would immediately reveal a spurious signal within $n_{\mathrm{OMS}}(t)$, if it was large enough to be detected against the noisier GRS data. Indeed, the feedback loop is driven by the OMS readout (see Fig. 2.16). If it was faulty and included spurious transients, the control loop would try to counteract a non-existent force, leading to a real TM2 displacement. Moreover, TM2 would be subject to additional gradient forces.

In particular, to yield the detected shape in Δg, an OMS glitch signal $\ddot{n}_{\mathrm{OMS}}(t) + \omega_2^2 n_{\mathrm{OMS}}(t)$, should follow one of the templates in Eqs. (7.1, 7.2, 7.5). Defining $h(s)$ as the Laplace transform of the template in Δg, then the Laplace transform of the associated glitch in $n_{\mathrm{OMS}}(t)$ would be given by $h(s)/\left(s^2+\omega_2^2\right)$.

As $h(s)$ is a rational function of s, the term $n_{\mathrm{OMS}}(t)$ would carry a diverging term $\propto \exp(+\sqrt{-\omega_2^2}t) \simeq \exp(+t/1.5\,\mathrm{ks})$, which after a few thousand seconds would dominate the data.

We depict this, for one of the glitches, in Fig. 7.16. The glitch is clearly visible in both $\Delta x_{12,\mathrm{GRS}}(t)$ and $\Delta x_{12,\mathrm{OMS}}(t)$, while it is not in their difference. In the right panel, we also show the inverse Laplace transform of $h(s)/\left(s^2+\omega_2^2\right)$, i.e. the signal one would observe in $\Delta x_{12,\mathrm{OMS}}(t) - \Delta x_{12,\mathrm{GRS}}(t)$, if the glitch were due to a spurious signal in $n_{\mathrm{OMS}}(t)$. The picture clearly shows that this source for this glitch is ruled out.

We have calculated the hypothetical signal in $n_{\mathrm{OMS}}(t)$, that would show up as the detected events, both in the ordinary and cold runs. We have found that such signals would have been clearly visible in the data if a sufficiently long observation time were available after the glitch occurrence time (up to 10^4 s for the weakest glitches). We

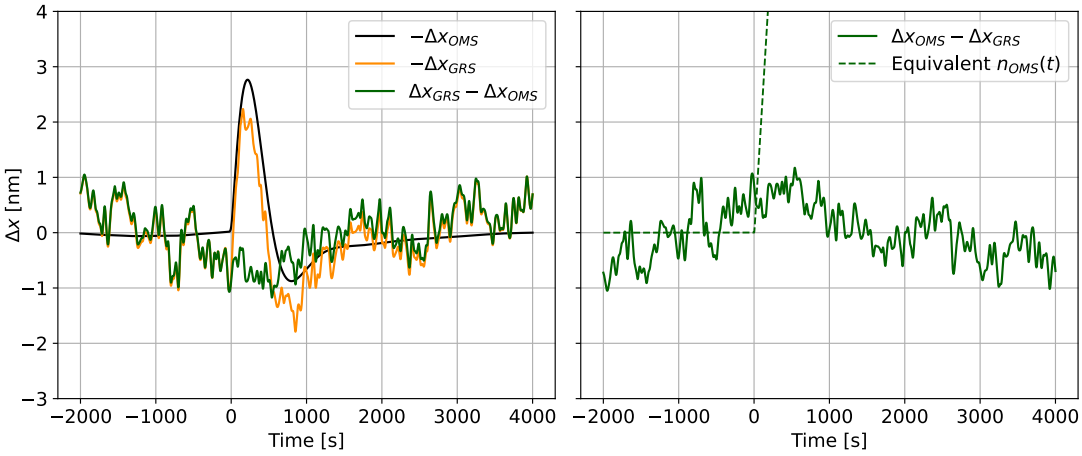


Figure 7.16: On the left panel, TM displacement $\Delta x(t)$, as measured by OMS (black), the GRS (orange), and their difference (green), for the short-lasting strong glitch of Dec16B. Despite the glitch lasting $\sim 8\,\mathrm{s}$, the $\Delta x(t)$ profile is much longer, due to the characteristic $<$ mHz frequencies of the actuation control loop. On the right panel, the difference $\Delta x_{12,\mathrm{OMS}}(t) - \Delta x_{12,\mathrm{GRS}}(t)$ is compared to the signal (dashed line) one would observe if the glitch was a spurious signal in $n_{\mathrm{OMS}}(t)$. Glitch parameters: $\Delta v = (22.2 \pm 0.01)\,\mathrm{pm/s}$, $\Delta = (7.82 \pm 0.01)\,\mathrm{s}$.

have found none in the actual data. Only for a few events (in cold runs) we could not reach any conclusion due to lack of sufficient data after the glitch.

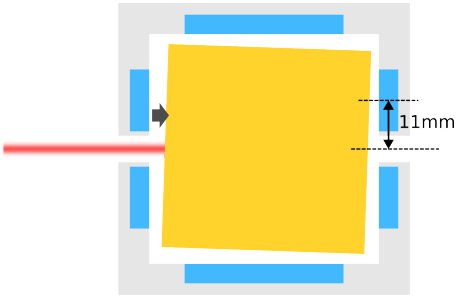
Concluding, impulse-carrying glitches are due to true forces acting on the test masses, and are not artifacts due to motion readout.

Note on FLGs. Low-impulse glitches consist mostly of features in $\Delta x_{12,\mathrm{OMS}}$. Moreover, they are not visible within the feedback force time series because of their spectral shape, having no power at low frequencies. However, they could still be either true acceleration spikes, or features in the interferometer output.

As already mentioned, glitches of this kind consist of a step, or a few-points outlier in $\Delta x_{12,\mathrm{OMS}}$, or some variation of those. In the case they were just features in the interferometer, $n_{\mathrm{OMS}}(t) - n_{\mathrm{GRS}}(t)$ would contain similar features (neglecting $\omega_2^2 n_{\mathrm{OMS}}$ for these fast signals). Unfortunately, the largest steps are a few tens pm high, while the resolution on step detection in the $\Delta x_{12,\mathrm{OMS}}(t) - \Delta x_{12,\mathrm{GRS}}(t)$ is not better than $\sim 1\,\mathrm{nm}$. Thus, we were not able to discriminate FLGs, nor exceptions, between true forces and interferometer readout errors. More details about the nature of FLGs will be provided in Sec. 7.4.1.

7.3.2 Associated differential torque

After proving that LPF ICG glitches are forces acting on the TMs, we investigate the presence of similar events in the angular acceleration time series.
If ICGs were well-localized forces, with a force vector not passing through the TM center of mass, they would induce a net torque, which would be detected in the out-of-loop angular acceleration time series.



As schematized on the right, if the force was directed in the x direction, with an application point different from the laser spot, the event would have a lever arm along the z or y directions (inducing, respectively, a torque about ϕ or η). The same would hold even in case the force was not well-localized, but rather distributed: a total force and torque could anyway be identified. This would be the case of a force originating from the actuation electrodes, which would show up with a characteristic lever arm $|r_\phi| = 11\,\mathrm{mm}$ [78].

As defined in Sec. 2.2.5, in close analogy with what we did with $\Delta g(t)$, we measured the differential out-of-the-loop torque per unit moment of inertia on the test masses, both around the y (η) and around the z (ϕ) axes, resulting in the $\Delta\gamma_\eta(t)$ and $\Delta\gamma_\phi(t)$ time series. As the rotational motion of the spacecraft is rather intense, it completely dominates the common-mode angular acceleration of the TMs. In addition to just analyzing the differential time series, we recalibrated torques and angular rotations to maximize the rejection of such a large common mode disturbance.

In the hypothetic case of non-centered force, the transient showing up in the angular acceleration time series would have the same time profile as in the linear acceleration. For each of the ICGs within the $\Delta g(t)$ time series, we have fitted both the $\Delta\gamma_\phi(t)$ and the $\Delta\gamma_\eta(t)$ time series to exactly the same template in Eqs. (7.1, 7.2, 7.5), with same time of occurrence and same time parameters, leaving the amplitude $\Delta\Omega_{\phi,(\eta)}$, as its sole fitting parameter. This parameter can be identified as an effective variation of angular velocity.

Fitting technique. On a data stretch including an ICG glitch, we perform a least squares fit with the glitch templates $h_1(t)$ or $h_2(t)$, with fixed τ_1, τ_2, but variable amplitude. We perform a sliding fit, varying t_0, aiming at identifying the glitch occurrence with a clear amplitude ($\Delta\Omega_{\phi,(\eta)}$) peak. We modify the least square fit to subtract a 2nd-order polynomial from the time series, to subtract the background low-frequency noise. This is needed to remove as much as possible the red background noise.

For the single-τ model (analogously for the two-τ case) we generate the templates $f_j(t)$, as

$$f_1(t) = t/\tau^2\, e^{-(t-t_0)/\tau}; \quad f_2(t) = 1; \quad f_3(t) = (t-t_0); \quad f_4(t) = (t-t_0)^2 \qquad (7.11)$$

The standard least square method allows to find the coefficients β_j, such that data $y(t)$ are well approximated by $\sum_j \beta_j f_j(t)$. The well-known result is $\boldsymbol{\beta} = (\boldsymbol{F}^T\boldsymbol{F})^{-1}\boldsymbol{F}^T\boldsymbol{Y}$, where the capitalized variables indicate the discrete-time representations of $f_j(t)$ and $y(t)$. We apply this method on consecutive windows at varying t_0, performing the convolution

$$\boldsymbol{\beta}(\delta t) = \left(\boldsymbol{F}^T\boldsymbol{F}\right)^{-1}\left(\boldsymbol{F}^T * \boldsymbol{Y}\right)(\delta t), \qquad (7.12)$$

where $\boldsymbol{Y}$ is a data stretch containing the glitch, which is longer than the template

$\boldsymbol{F}$ to allow convolution. The first parameter β_1, at zero delay $\delta t = 0$, represents the estimate of the angular velocity variation $\Delta\Omega_{\phi,(\eta)}$, as shown in Fig. 7.17. In case of correspondence, a marked peak at $\delta t = 0$ should be visible.

We estimate the uncertainty on $\Delta\Omega_{\phi,(\eta)}$ by repeating the above fitting/convolution procedure, over the stretch $\Delta\gamma_{\phi,(\eta)}(t_1 + \delta t, t_2 + \delta t)$, with t_1 and t_2 the time bounds of the actual stretch that contains the glitch. We thus generated a series $\Delta\Omega_{\phi,(\eta)}(\delta t)$, with $|\delta t| \leq 20(t_2 - t_1)$, which shows, at $\delta t = 0$, a marked peak above the background jitter, whenever the torque is significant (see Fig. 7.17).

As the $\Delta\Omega_{\phi,(\eta)}(\delta t)$ time series has an intrinsic autocorrelation over a scale $\delta t \simeq t_2 - t_1$, it contains in practice about 40 independent data points. Thus, for error estimation, we use the peak absolute value $\delta\Omega_{\phi,(\eta)}$ of the series, calculated on the data outside the central stretch $|\delta t| \leq t_2 - t_1$. More specifically, if the central peak does not exceed, in absolute value, $\delta\Omega_{\phi,(\eta)}$, we take $|\Delta\Omega_{\phi,(\eta)}| \leq \delta\Omega_{\phi,(\eta)}$.

If, contrarily, the peak exceeds in absolute value $\delta\Omega_{\phi,(\eta)}$, we take this as the error. It must be noted though that, for Gaussian statistics, the maximum absolute value among 40 independent samples falls in the interval $(2.4^{+0.5}_{-0.4})\sigma$. Thus the confidence interval associated with such an error is greater than 95%. The result of this estimation can be seen in Fig. 7.18, though it requires some more parameters to be understood.

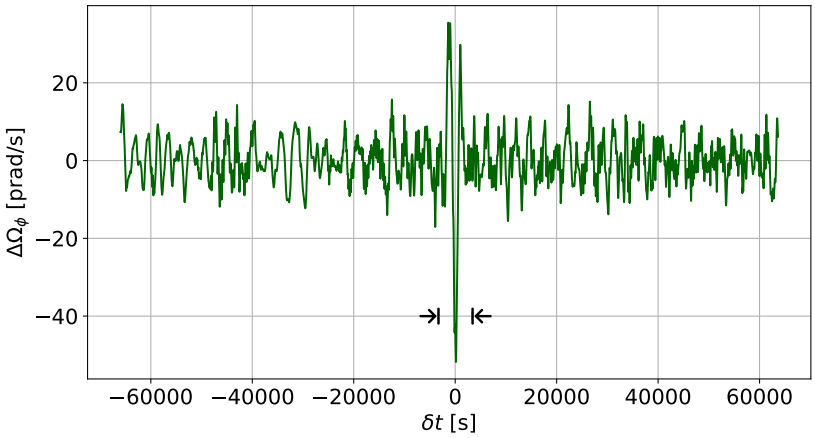


Figure 7.17: An example of the $\Delta\Omega_\phi(\delta t)$ series for one of the very few glitches with significant associated torque. The value in $\delta t = 0$ is the result of the fit for native data. The arrows indicate the interval excluded for error evaluation.

Fitting results. To understand better the properties of glitches in this analysis, we introduce an effective *lever arm* defined as

$$r_{\phi,(\eta)} = \frac{I_{zz,(yy)}}{m}\frac{\Delta\gamma_{\phi,(\eta)}}{\Delta g} = \frac{I_{zz,(yy)}}{m}\frac{\Delta\Omega_{\phi,(\eta)}}{\Delta v} \tag{7.13}$$

For a single force, applied normally to an x-face of either test mass, r_ϕ would be the distance between the force application point and the vertical symmetry axis of the face. A similar interpretation holds for r_η. For a point-like force, given the size and shape of the TM, the maximum value for both the effective arms $|r_\phi|$ and $|r_\eta|$ would be 23 mm, while there would be no upper limit in case of a distribution of forces.

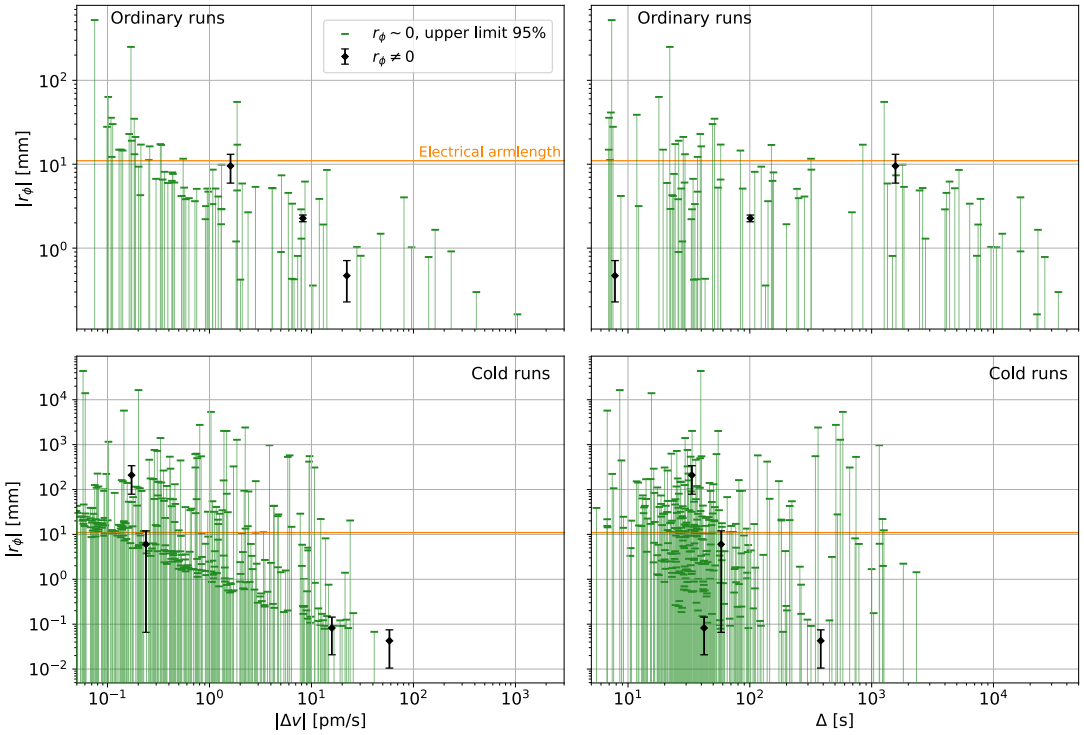


Figure 7.18: Absolute lever arm $|r_\phi|$, as defined in Eq. (7.13), as a function of the absolute impulse $|\Delta v|$ (left) and duration Δ (right) for impulse-carrying glitches. Upper panels refer to ordinary runs; lower panels, to cold runs.
The black data points represent the glitches for which the lever arm has been found to be significantly different from zero. The green horizontal segments represent the upper bound for $|r_\phi|$, in case the estimate of the transferred angular velocity is compatible with zero. For 9 of the 98 impulse-carrying glitches of the ordinary runs, we were not able to perform the analysis on the $\Delta\gamma_\phi(t)$ data series. The orange horizontal line refers to the 11 mm "electrical" reference, corresponding to the characteristic lever arm of electrical single-electrode forces.

The analysis results for r_ϕ in ordinary runs are shown in Fig. 7.18. We find a few events with any of the lever arms significantly different from zero, which we provide in Table 7.2.

We make some observations on the results in Figure 7.18:

1. As already expected from Eq. (7.13), the error on the effective arm r_ϕ decays as $|\Delta v|^{-1}$, decreasing below about 1 cm for $|\Delta v| \gtrsim 2$ pm/s. Increasing the glitch impulse increases the precision of the lever arm estimate.

2. Only 3 of the 98 impulse-carrying glitches in ordinary runs show an effective armlength that is significantly different from zero. In 62% of the cases, glitches have an armlength compatible with zero, and at the same time there is sufficient resolution to exclude the hypothesis of an electrode origin to the glitch. Specifically, we exclude the electrical-origin hypothesis if the "upper limit" uncertainty is less than 7 mm (or 2/3 of the characteristic "electrode arm" of 11 mm). If any of these relatively high SNR glitches had had an 11 mm effective armlength, they would have been detected with peaks more than 50% above our background in

Fig. 7.17. This fraction decreases to 50% for glitches with $\Delta \leq 1\,\mathrm{min}$.

3. The time series $\Delta\gamma_\eta(t)$ is significantly noisier than $\Delta\gamma_\phi(t)$. This gives larger errors on r_η than on r_ϕ, and makes it less likely to find lever arms significantly different from zero. Just for only about 32% of the glitches the error upper bound is smaller than (2/3) 11 mm. However this figure has no physical meaning for η, as the y-face electrodes have a different configuration.

4. The angular time series $\Delta\gamma_{\phi,(\eta)}(t)$ are complex to analyze, as they contain multiple fast events and other non-stationary features. This makes the results of our search procedure noisier, reducing, for instance, the fraction of glitches, for which the upper error bound to r_ϕ is less than $(2/3)\,11\,\mathrm{mm}$, to about 45%, and to 39% if $\Delta \leq 1\,\mathrm{min}$, as shown in Fig. 7.18, in the lower panels.

5. Moreover, we note that, of the 3 zero-crossing glitches transferring impulse in Table 7.1, none of them is associated with a measurable torque, measured with the same analysis as in this section. The error on r_ϕ associated with the stronger ones (one in ordinary runs and one in cold runs), is lower than (2/3) 11 mm.

   As a final note, in addition to the glitches listed in Table 7.2, our procedure also finds 3 glitches in the cold runs for which the peak in $\Delta\Omega_{\phi,(\eta)}(t)$ is significantly displaced from the true time of occurrence. We believe that these events are due to the accidental coincidence between a force glitch and an unrelated feature in the torque data.

Table 7.2: Glitches with any of the lever arms significantly different from zero. Figures above the horizontal line refer to ordinary runs, figures below the same line to cold runs. Errors correspond to confidence > 95% (see text).

Δv [pm/s]	Δ [s]	r_ϕ [mm]	r_η [mm]
-8.23 ± 0.03	101.1 ± 0.4	-2.3 ± 0.2	...
1.6 ± 0.2	1560 ± 140	-9.5 ± 3.5	...
22.22 ± 0.01	7.82 ± 0.01	0.5 ± 0.2	-1.2 ± 0.4
(*) 58.10 ± 0.05	773.1 ± 0.4	0.04 ± 0.03	0.12 ± 0.11
(*) 15.85 ± 0.01	78.1 ± 0.2	0.08 ± 0.06	–
(*) 0.17 ± 0.01	61 ± 7	-210 ± 130	-260 ± 170
-25.68 ± 0.02	2100 ± 5	-1.2 ± 0.7	...

* Likely to be accidental coincidences between unrelated events.

7.3.3 Joint GRS analysis along y and z axes

We inspected the time series of the forces on the test masses along directions orthogonal to x, i.e., y and z, defined as Δg in Eq. (2.6), but with GRS readouts. Nevertheless, the differential force sensitivity along these axes is too low, with respect to that along x, to give any significant information. First, motion along y and z is noisier than

along x due to the lack of an interferometric readout. Second, actuation forces along those degrees of freedom were applied to follow the spacecraft's rotational motion. Indeed, the spacecraft's rotational motion was dominated by yaw and pitch jitters, causing, respectively, differential forces on the TMs along y and z.

Even though it is in principle possible to subtract such inertial forces decorrelating the common-mode applied torque along η and ϕ, it is practically very hard to completely remove such a big effect. Figure 7.19 in the left panel shows the PSDs of $\Delta g_{x,\mathrm{GRS}}$, $\Delta g_{y,\mathrm{GRS}}$, $\Delta g_{z,\mathrm{GRS}}$, as compared to the non-deglitched Δg_{ND} and the deglitched Δg, in run Apr16A. In the right panel, we show that no features are visible in coincidence with the loud long-lasting event.

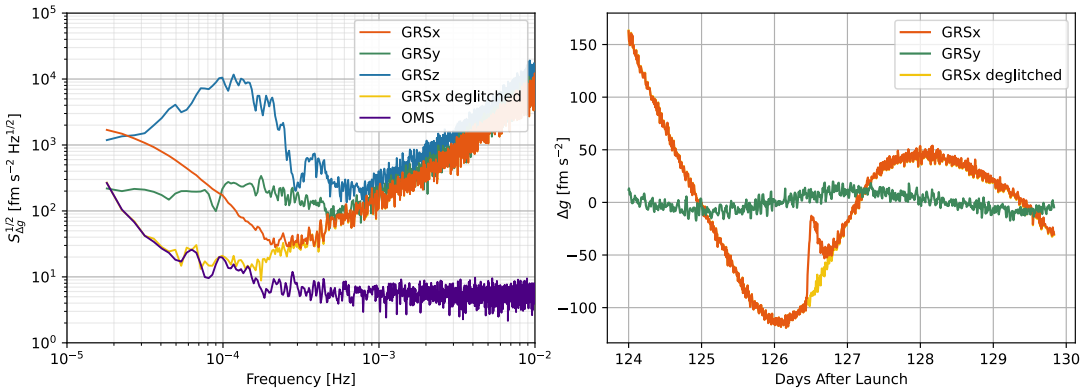


Figure 7.19: (left) Comparison between the out-of-loop differential accelerations $\Delta g_{x,\mathrm{GRS}}$, $\Delta g_{y,\mathrm{GRS}}$, $\Delta g_{z,\mathrm{GRS}}$, defined analogously to Δg in Eq. (2.6), in run Apr16A. The label ND indicates non-deglitched and deglitched data. Note that run Apr16A contains the loudest hour-lasting glitch, whose impact is visible in the non-deglitched x data. OMS in the legend indicates the standard Δg.
(right) Lowpassed $\Delta g_{x,\mathrm{GRS}}$ and $\Delta g_{y,\mathrm{GRS}}$ data. The loud ICG event is well visible in the $\Delta g_{x,\mathrm{GRS}}$ data, and an event with similar amplitude would be well visible in $\Delta g_{y,\mathrm{GRS}}$. Nevertheless, there is no trace of events.

We found indeed no evidence of coincidences between ICG events and transients in y and/or z, nor evidence of non-coincident events, showing the typical ICG templates in Eqs. (7.1, 7.2, 7.5). As said, this does not exclude the presence of ICG transients along such axes, especially regarding long-lasting events, for which the detection threshold was much higher.

7.4 Joint analysis with other time series

7.4.1 Interferometer channels: low-impulse glitches

Contrarily to ICGs, we find a significant number of coincidences between the FLG glitches and fast transients in the interferometer channels, other than $x_{12,\mathrm{OMS}}$. We find coincidences with events in the measurement channel $x_{1,\mathrm{OMS}}$, the reference, and the frequency channels (R, F, see Sec. 2.1.2 for details). The total number of glitches showing a coincidence in the interferometric channels is 81/152. Some of them also show multiple coincidences. However, due to the complexity of the data of the cold

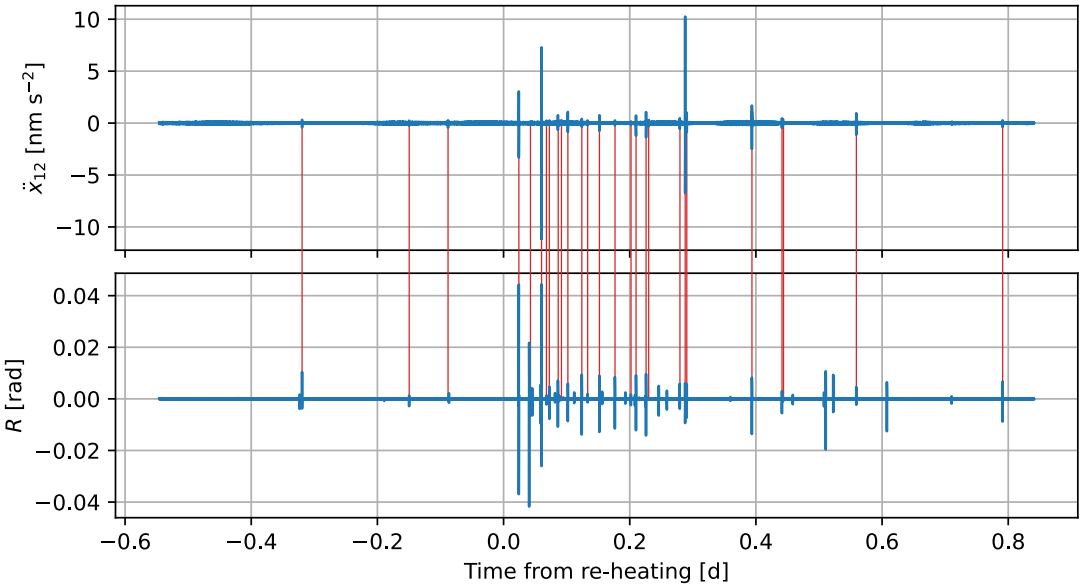


Figure 7.20: Correspondence between events in the $x_{12,\mathrm{OMS}}$ time series (upper panel) and the R interferometer channel (lower panel), in the last part of run May17D, after re-heating. Red lines mark FLGs. Figure 7.21 provides a close-up of two adjacent FLGs, showing the time structure in the R channel.

time series, it is rather likely that some events and some coincidences have gone undetected. This does not change our conclusions, nor alter our statistics.

Interferometer reference/frequency channels

Out of 152 low-impulse glitches detected in cold runs, 55 correspond to events in the R channel, and 4 correspond to events in the F channel. In the latest part of the cold runs, well visible in Fig. 7.10 on page 143, and again in Fig. 7.20, tens of glitches of this kind appear with non-Poissonian arrival time after the heaters were turned on to bring the temperature back to standard operating conditions. Those events showed a rapid deviation in the reference channel, as big as some milliradians. In Fig. 7.20, we show the coincidences between Δg FLGs and events in the R interferometer, in the latest part of run May17D. All these events showed in the R channel a characteristic shape lasting some tens of seconds, due to the internal control loops of the optical interferometer system [34]. It can be noticed in the close-up in Fig. 7.21. The shape in either the reference or frequency channels is different from that of Δg, which is much faster.

We recall that the readout of the R channel is inherently subtracted from the readouts of the other channels, including $x_{12,\mathrm{OMS}}$. The subtraction was performed onboard at a 100 Hz rate, before downsampling and downlinking at the nominal sampling frequency. We note that the transient detected in $x_{12,\mathrm{OMS}}$, as can be seen in Fig. 7.21, could also be present in the original $x_{12,\mathrm{OMS}}$ time series, and the features that we detect could just be the result of a non-perfect data subtraction. Indeed, a ~ 5 mrad variation, as detected in R, would correspond to a spurious ~ 0.5 nm variation in $x_{12,\mathrm{OMS}}$, which is of the order of the detected transients.

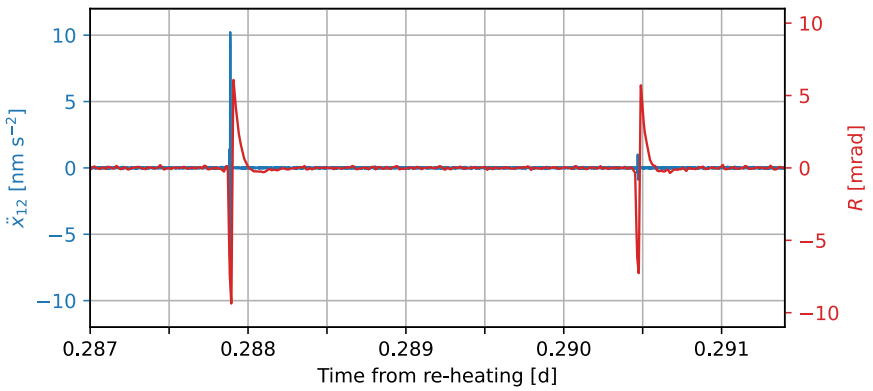


Figure 7.21: Close-up of two adjacent FLGs, showing a 400 s timespan from Fig. 7.20, in run May17D.

Interferometer x_1 channel

We find that some of the low impulse glitches belonging to cold runs coincide with fast events in the $x_{1,\mathrm{OMS}}$ interferometer, showing up as fast features in $\ddot{x}_{1,\mathrm{OMS}}(t)$. Among the 152 two-sided glitches detected in cold runs, 28 showed counterparts in $\ddot{x}_{1,\mathrm{OMS}}(t)$. Most of these events were found during the cold run May17B, with most of the times a time profile closely resembling that in $\ddot{x}_{12,\mathrm{OMS}}(t)$, as can be seen in Fig. 7.22. The time profile resembles that in $\ddot{x}_{12,\mathrm{OMS}}(t)$, apparently with a sign inversion and a different amplitude. However, the amplitude ratio does not seem to follow a predictable template, hence we do not analyze it further.

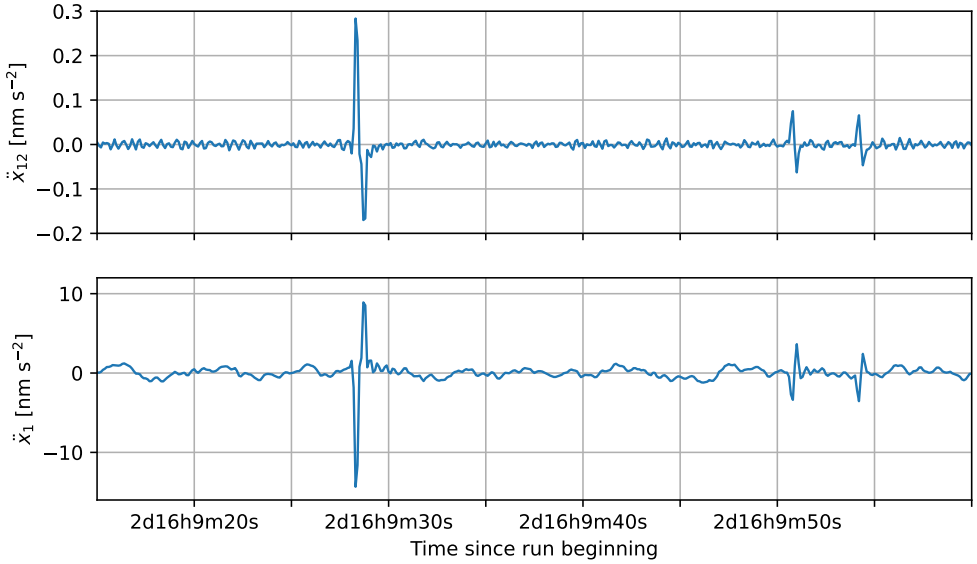


Figure 7.22: Unfiltered portion of the cold run May17B, showing FLG events in $\ddot{x}_{12,\mathrm{OMS}}(t)$, along with corresponding events occurring in $\ddot{x}_{1,\mathrm{OMS}}(t)$.

7.4.2 Interferometer channels: impulse-carrying glitches

No coincidences are found between interferometer channels and ICGs following templates in Eqs. (7.1, 7.2, 7.5). The only coincidence that we find between interferometer channels and impulse-carrying glitches occurs between the zero-crossing glitch on the second line of Table 7.1 and a spike in the $\ddot{x}_{1,\mathrm{OMS}}(t)$ time series, which is represented in Fig. 7.5 on page 139. The spike had no detectable counterpart in the time series of the force commanded onto the spacecraft by the drag-free control loop, as expected for such a fast feature.

In analogy with the case of the differential measurement, a spike like this might be either due to a spike in $g_1 - G$, with g_1 the force per unit mass on TM1 along x, and G that on the spacecraft, or to a feature in the $x_{1,\mathrm{OMS}}$ interferometer readout. When interpreted as a force, its total impulse, $\Delta v \simeq 0.8\,\mathrm{nm/s}$, is much larger than the corresponding impulse, Δg ($\Delta v \sim 0.8\,\mathrm{pm/s}$, in Table 7.1). This rules out that the spike is in g_1, and leaves only the options that it might have been due to a force impulse on the spacecraft, or to a feature in the $x_{1,\mathrm{OMS}}$ interferometer readout.

7.4.3 Other time series

We analyzed the time series of magnetometers and of the various thermometers of the LTP, without finding any coincidences. In the discussion, Sec. 8.3, we analyze the magnetic and thermal effects that could potentially induce force on the TM. We combine sensor sensitivity and observations, to evaluate the likelihood of glitches being of magnetic/thermal origin.

We also inspected the time series of the inertial forces, calculated as explained in [2], and found no coincidence. Note that these series have been subtracted from the observed acceleration to form Δg, in the post-processing pipeline. Thus, a true force glitch in the inertial forces should have emerged from the original acceleration series and should have disappeared after subtraction. Our check was meant to exclude the presence of non-force data artifacts in the inertial force series that may have been transferred to Δg because of miscalibration in the subtraction, which would have left a residual glitch in Δg.

7.5 Search of LPF-like ICG glitches in torsion pendulum data

There is no ultimate proof regarding the physical origin of LPF-like ICG glitches. However, the one certainty is that ICGs are caused by real forces, acting likely on one TM at a time. One could indeed wonder if those events could be detected on ground, analyzing data of already existent experiments, or designing new ones for that purpose. For instance, one of our initial questions was if LPF glitches could be detected in torsion pendulum facilities at the University of Trento. In this section, slightly off-track, we analyze the detectability.

The University of Trento hosts two Torsion pendulum (TP) facilities, designed to perform on-ground testings on GRS engineering and qualification models, allowing for on-ground preparation and analysis of the LPF and LISA subsystems. The TPs employ, at the moment of writing, EH GRS models closely resembling the one flown on LPF; in the near future, they will host LISA-like GRS models. The sensitivity to acceleration of such systems is not as high as the LPF one, especially at low frequencies. Nevertheless, the high fidelity and representativeness of the flight instruments makes it interesting to investigate on the presence of LPF-like glitch transients in those data.

1TM pendulum. Two systems have been in use. The one-test-mass (1TM) torsion pendulum [80] was the first one to be built. It has been operational in the first phase until 2015, and it is now being recovered. In its latest configuration [79], the experimental setup used a representative LPF-like electrode housing prototype, using flight-model gaps to the inner TM. The EH surrounded a hollow Au-coated $(46\,\mathrm{mm})^3$ TM, weighing 100 g, suspended via a 40 µm fused silica fiber, with high-quality factor $Q = (7.4 \pm 0.7) \times 10^5$, and very low torque noise. The name of the apparatus is easily understandable, since it only included a single TM.

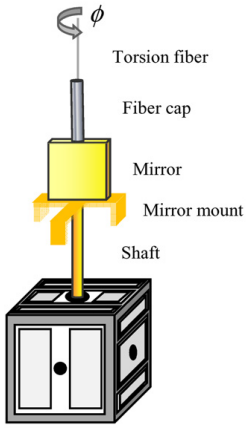


Figure 7.23: Scheme of the 1TM torsion pendulum at the University of Trento, in its latest configuration. Figure from [79].

Given the system configuration, it reached maximum sensitivity about the torsional degree of freedom about the vertical axis, indicated as ϕ in the picture on the right. Motion about ϕ was measured independently by two systems: the GRS itself via its capacitive sensors, and an autocollimator, much more sensitive, which used the mirror depicted in the figure on the right. An illustrative PSD is shown in Fig. 7.25. For our purposes, it is important to note that the 1TMTP could not be that sensitive along any other degree of freedom, thus excluding the direct measurement of translational forces.

Inspection of torsional data of the 1TMTP, upon conversion to torque time series as explained later, revealed a quite stationary behavior, showing no relevant transients, even upon filtering.

4TM pendulum. Another testing apparatus has been built at the University of Trento [81, 82], employing four LPF-resembling hollow test masses, mounted at the four ends of a 20 cm cross-shaped structure. Two opposite test masses are included in a LPF GRS prototype, closely resembling the flight model. It is clear that such a pendulum is, as well as the 1TM, only sensitive to torques about the vertical axis. However, in this configuration, the torsional motion also includes translational forces on the TMs, hence providing a measurement of linear forces.

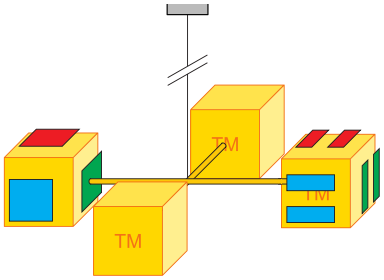


Figure 7.24: Schematic of the 4TM torsion pendulum. Figure from [81].

The used configuration, however, reduces sensitivity: to hold the 460 g suspended element, a 50 µm tungsten fiber has been used, at the expense of increased thermal noise and a lowered sensitivity. As well as the 1TM model, the 4TM pendulum employs two independent measurements of the pendulum angle, namely an autocollimator and the capacitive GRS itself.

We note, for our purposes, that the 4TM pendulum is, contrarily to the 1TM model, sensitive to linear forces, and thus possibly sensitive to LPF-like ICG transients. However, we also notice that the 4TM pendulum is much more subject to many kinds of transients, often showing spikes and spurious events.

Here we describe the method that we used to analyze the torsion pendulum time series, looking for LPF-like ICG events. The two measurement sensors (autocollimator and GRS) are used in coincidence to give an estimate of a joint SNR, that takes into account auto- and cross-noise correlations. In this way, spurious readout noises happening on only one of the two sensors are efficiently discarded. Here we anticipate that, in the 7 noise runs analyzed, no SNR exceeding 4σ is found, with no clues of LPF-like glitches.

7.5.1 Torque noise PSD evaluation

First of all, torque time series have to be computed from the angle time series. For consistency, we convert torque into the equivalent LPF differential acceleration.
The torsion angle ϕ is associated with the z-axis torque N_ϕ through the differential equation

$$N_\phi(t) = I_0 \left[\ddot{\phi}(t) + \frac{\omega_0}{Q}\dot{\phi}(t) + \omega_0^2\phi(t)\right] \quad \Rightarrow \quad N_\phi(\omega) = I_0 \left(-\omega^2 + \frac{i\omega\omega_0}{Q} + \omega_0^2\right)\phi(\omega) \tag{7.14}$$

where I_0 is the moment of inertia, $\omega_0 = 2\pi/T_0$ the proper angular frequency, associated with the period T_0, and Q the quality factor. Torque is then converted to the equivalent single-mass LPF acceleration, via the lever arm r, using $g = 2N/(rM)$, where $M = 1.928\,\mathrm{kg}$. In the case of the 4TM pendulum, r is known; conversely, in the 1TM case, we could just assume it.

In Fig. 7.25, we show the PSDs of the equivalent LPF acceleration in both the 1TM and 4TM runs, with autocollimator and capacitive measurements. For the 1TM

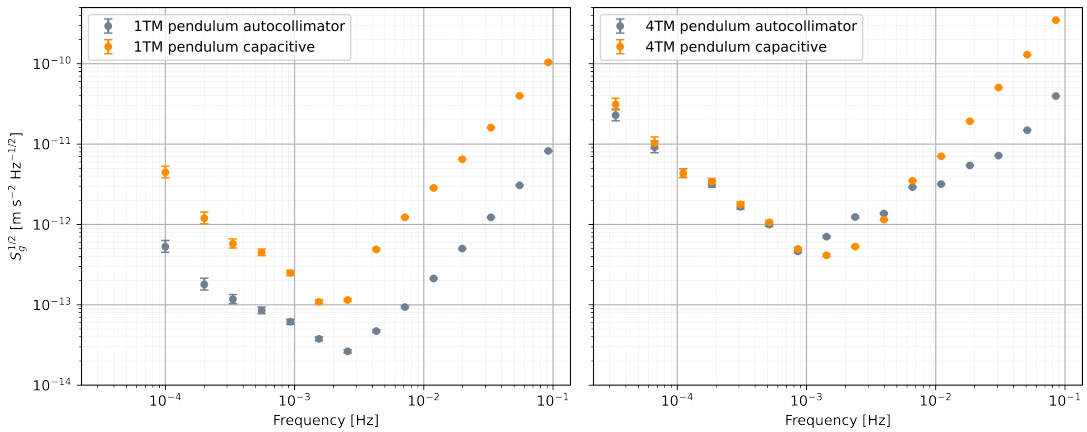


Figure 7.25: PSDs of the equivalent single-mass LPF acceleration noise, for the 1TM pendulum (left panel, courtesy William J. Weber), and the 4TM one (right panel, courtesy Davide Dal Bosco), autocollimator and capacitive data. Frequencies in the 4TM case are lower because the analyzed data stretch was longer. For the 1TM case, we assume a lever arm of 1 cm when converting to equivalent LPF acceleration.

case, we assume a lever arm of 1 cm.

7.5.2 Matched-filter application to torsion pendulum data

The standard matched-filtering technique, also known as the Wiener filter, allows to define an optimal filter for template detection, based on the background noise PSD of the measurement instrument. Based on the transient signal template, the filter allows to extract the amplitude of the signal. In the basic case, a single detector is used, and the template is searched in its time series. In a more advanced scenario, such as the LIGO one, two completely different detectors are used to detect the same GW transient, slightly time-shifted); in that case, each of the detectors has its own background PSD noise [83].

The torsion pendulum case is different. Two detectors are in use (autocollimator and capacitive sensor), which are measuring the same quantity, hence in the presence of partially correlated noises. Readout noises differ between the two systems, but real force noise should in principle be the same. A well-defined signal-to-noise ratio post-processing should discard events happening in just one of the two sensors, and should peak in presence of joint template-matching events. For the sake of generalization, we look at the problem from a generic point of view.

We define such a matched-filter in App. C.2, and apply it to pendulum data. We find the frequency-domain version of the filter template, matching the single-τ glitch shape, Eq. (7.1), with some typical values of τ. A systematic search would include an exploration of the τ parameter space, and an analysis of the multi-τ case, which we do not perform for detectability reasons, which will be clearer later. Rather, we just analyze some values of τ.

First, in Fig. 7.26, we show the two-detector detectability thresholds for the 1TM and the 4TM pendulums, compared to the detected LPF ICG glitches. We show thresholds at SNR = 5, since it would be inappropriate to claim detection at lower

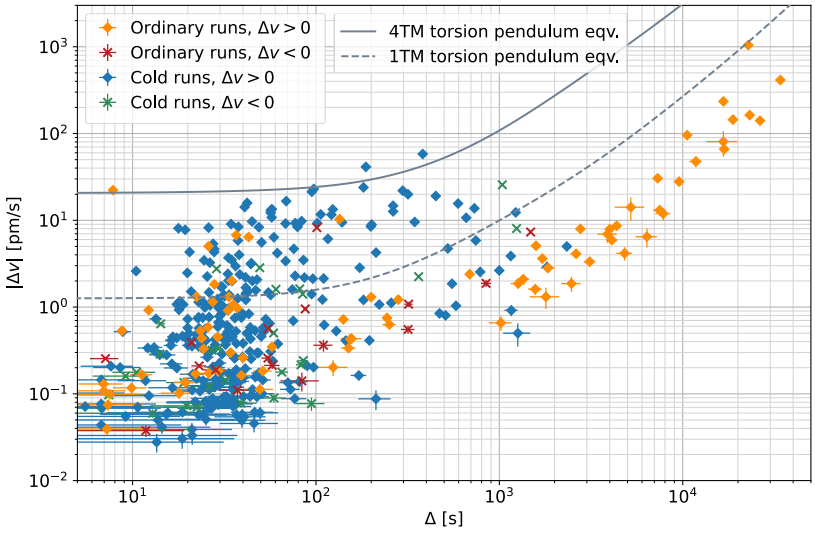


Figure 7.26: Detectability threshold of ICG glitches matching the template in Eq. (7.1), in the case of coincident two-detector search in torsion pendulum data. The grey lines represent SNR = 5 lower thresholds, for (solid) the 4TM pendulum and (dashed) the 1TM one. Grey lines represent the equivalent LPF linear acceleration, corresponding to the measured TP torque. In the 1TM case, the lever arm is assumed to be 1 cm; if it was 1 mm, the two pendulums would basically have the same resolution.

values, given the presence of other species of glitches in TP data.

We draw some conclusions on the detection of LPF-like events in torsion pendulum data.

1. It is highly unlikely to find LPF-like ICG glitches with the 4TM pendulum apparatus, without implementing mechanisms to enhance the glitch impulses. Long-lasting glitches (hours) could not be detected at a reasonable SNR level, and even the short-lasting ones would be difficult to identify. However, if the glitch impulse could be enhanced by a factor ~ 10, glitches detected during cold runs, lasting more than 100 s, could be well identified.

2. Regarding the 1TM pendulum, it could hardly detect long-lasting glitches detected during the LPF ordinary runs. It could however detect short-lasting events, especially those found during the cold runs. It is however important to remark that the 1TM system, as is, does not provide a measurement of linear forces, and thus measures a different quantity. LPF ICG events showed basically no torque, and the small torque shown by some of them could not be identified with the 1TM apparatus. The threshold provided in Fig. 7.26 highly relies on the chosen fictitious lever arm, 1 cm. If it was 10 times smaller, the threshold would be 10 times higher, thus preventing the measurement of any LPF-like glitch.

3. For such reasons, we find it not necessary to run a systematic search on the entire parameter space. Rather, we proceed as follows:

   (a) We inspect the time series by eye, applying a 1000 s Blackman-Harris filter

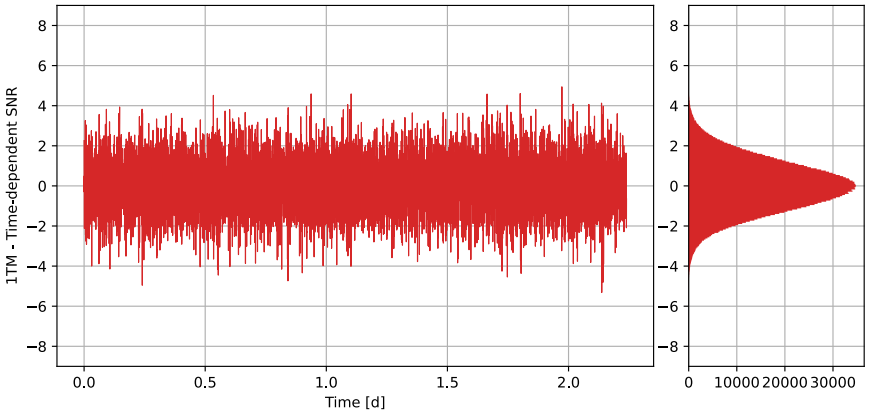


Figure 7.27: Illustrative run, 1TM data filtered with matched-filter, as defined in text. The filter matches the template in Eq. (7.1), with $\tau = 50\,\mathrm{s}$. On the right, a histogram of filtered data is shown. No matches at high SNR are found.

to detect features, to about 10 pendulum runs. We find no relevant events resembling the LPF ones, neither in the 1TM, nor in the 4TM data.

(b) We run our two-dimensional matched-filter to 1TM torsion pendulum data, looking for events with $|\mathrm{SNR}| \gtrsim 5$. We look for events matching the template in Eq. (7.1), varying τ. Figure 7.27 shows the SNR results with $\tau = 50\,\mathrm{s}$ in an illustrative run of the TM1 pendulum.

To conclude, we find no events matching the LPF template within the time series of the 1TM torsion pendulum, nor in the 4TM one. However, even if correspondence was found, it would hardly be conclusive: at present, suppressing the high-frequency noise in torsion pendulum data requires strong low-pass filters to detect LPF-like glitches. Even if an event was detected, it would be hard to guess its shape, as it would closely resemble the shape of the employed filter.

We find that stronger conclusions could be drawn if glitches of the long-lasting kind were detected (see Fig. 7.26), but we already showed that those could hardly be detected at the moment. For future reference, aiming at detecting LPF-like events, a strategy to increase the total transferred impulse would be advisable and necessary.

8 Discussion and conclusions

This last chapter is dedicated to a discussion, and some possible explanations, about the results found in the presented analyses. The necessity of a discussion section arises from the fact that not all the observed features of LPF noise have been completely explained. The discussion is to some extent more qualitative than the previous analyses, as it is meant to propose possible mechanisms producing the observed features, which are not always precisely modelable. We analyze a series of mechanisms, one by one, and illustrate why such mechanisms may or may not fit observation. Finally, we give a last summary and conclusions.

As usual, we split the examination of the observed features into three parts,

1. In Sec. 8.1, we discuss the possible origin of the observed excess noise at low frequencies. We retrospectively look at the post-processing steps, especially those containing correction terms, and consider some additional unmodeled force-inducing effects. This includes a cross-over analysis with glitches, in which we look at the possibility that the excess noise results from undetected glitches.

2. In Sec. 8.2, we discuss the observations about pressure evolution in the VEs, and the LPF outgassing environment.

3. In Sec. 8.3, we thoroughly discuss the possible physical origin of the detected glitches, mostly dealing with the impulse-carrying ones, as they may majorly affect LISA's GW detection. In this perspective, we estimate the fraction of LPF glitches that would be detected by LISA, according to the current instrument requirements.

8.1 Discussion on the physical origin of the low-frequency excess

In Chapters 4 and 6, we presented a number of analyses about the sub-mHz noise branch, ultimately showing that there is a clear excess noise that does not have a clear source. We however found certainty about the nature of this observed noise. In Sec. 4.3, we indeed showed that the observed excess is caused by true forces disturbing the TMs free-fall motion. Chapter 6 was entirely dedicated to understanding the contribution of effects related to other LPF telemetries. We showed that, even though there is no smoking gun for processes generating such noise, there is a number of processes marginally contributing to it. For each of these processes, we estimated

the upper limits of their contributions. We showed that the dominant contribution at the first frequency is understood, while at higher frequencies (36 µHz–1 mHz) the contribution from the considered effects is no more than some percent.
Ultimately, as reported in Sec. 6.5, to our knowledge the strongest contribution comes from the actuation gain fluctuations. Such contribution, though, accounts only for $\sim 30\%$ above 0.1 mHz, and $\sim 10\%$ below it. This means that there is an unaccounted noise source, which is dominating the sub-mHz acceleration noise, that we still have not considered. We look at some more possibilities.
Another option, which clearly must be considered, is that some of the corrections applied in the post-processing to estimate Δg contained imperfections leading to a higher residual, i.e. either correcting less than necessary or even over-correcting.
We analyze these possibilities, one by one, starting by looking at the post-processing pipelines.

8.1.1 Post-processing imperfections

Looking back at Fig. 2.17, in the postprocessing pipeline the sub-mHz spectrum is dominated by commanded forces and non-inertial corrections. Gradient forces are orders of magnitude lower than the out-of-loop acceleration, as well as the in-loop acceleration, which is suppressed thanks to the actuation control. Hence, a hypothetical imperfection or inaccuracy in the estimates of the inertial forces might lead to an inexact result.

Centrifugal force estimation. The first correction step in the postprocessing pipeline is the subtraction of the centrifugal forces. As described in [1, 3], the contribution from such effect is calculated using a combination of the quasi-DC angular velocity from star-tracker telemetries, and the in-band applied torques. There are no free fit parameters, so in the end, the time series is exactly known. It is calculated as in Eq. (2.7). The magnitude of the correction is shown in Fig. 8.1. The black points represent the common excess noise, as computed in Fig. 4.5, which is representative of the excess to be explained. The brown curve is the centrifugal acceleration in run Feb17B, and the brown area represents the variability during the 13 selected `UURLA` runs, at 1σ.

Even though during some runs the spacecraft jitter was really intense, with centrifugal acceleration exceeding the common excess noise, it was not always the case. In Feb17B, for example, the correction only exceeds the excess in two/three points. Moreover, the spectral shape of the centrifugal correction does not resemble a f^{-2}-shaped PSD. Hence, an inaccuracy in the estimate of the centrifugal correction could not lead to an excess as the measured one. Moreover, it is very unlikely that this time series is affected by relevant inaccuracies, higher than a few %; as said, there were no free parameters in the computation.

Tangential force estimation. The second considered inertial force is the tangential Euler force. As shown in Eq. (2.9), its calculation is based substantially on the applied common-mode torque time series, whose error is negligible. However, the calculation also includes a dependence on two free parameters representing angular misalignments, which might be affected by uncertainty. The estimation of the contri-

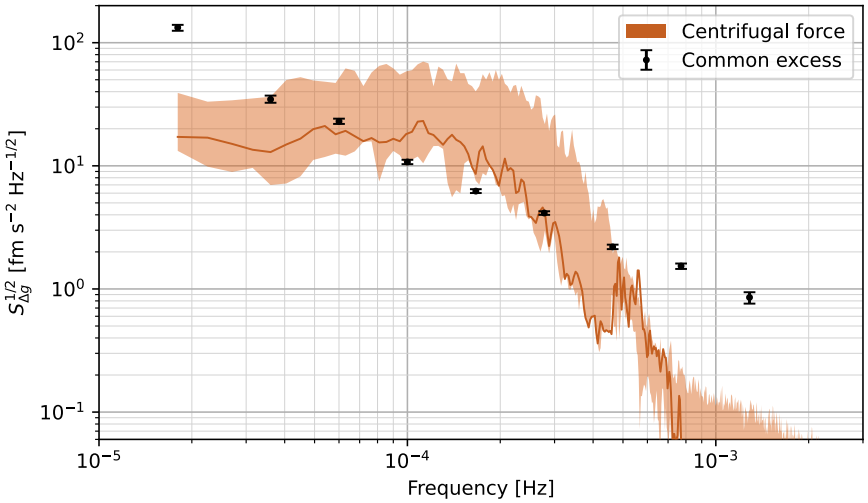


Figure 8.1: Comparison between the common excess noise (black points), the centrifugal acceleration in run Feb17B (brown solid line), and the variability during the 13 selected `UURLA` runs, as 1σ quantiles (brown area).

bution is shown in Fig. 8.2. Colors and lines have the same meaning as in the previous figure.

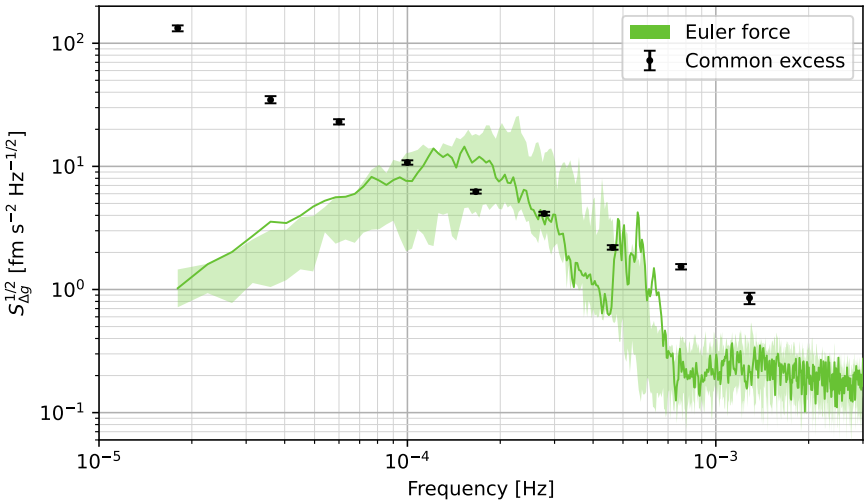


Figure 8.2: Comparison between the common excess noise (black points), the tangential Euler acceleration in run Feb17B (green solid line), and the variability during the 13 selected `UURLA` runs, as 1σ quantiles (green area).

Estimation of the two angular parameters was performed by splitting the mission into three epochs, corresponding to the three TM alignment conditions in Table A.2. Re-doing the fit in single runs, slight variations of these parameters can be noticed, of the order of a few %. Changing parameters may lead to slightly lower excess noises. Nevertheless, as noticeable in Fig. 8.2, Euler force only affects a few frequencies around 0.1 mHz, making it impossible to explain the observed noise below 0.1 mHz. Imperfections in the fits for the angular misalignments can not lead to significant

variations in that figure.

8.1.2 Gravitational noise

Possibly, force noise could originate from unmodeled sources of gravitational forces, induced by modifications of the gravitational field at the test mass locations. The modelization of such a phenomenon is however really complex, as many sources could be contributing to it, increasing the number of parameters to take into account. We can however make some observations, aiming at understanding which are the key points in force generation, and understanding to which extent we can rule out these mechanisms.
As it is computationally easier, we refer to $\boldsymbol{G}(\boldsymbol{x})$ as the gravitational field generated by the TM, of mass M. We assume that m is the mass of a point-like object, which moves by a small distance $\delta\boldsymbol{x}$ about its location $\boldsymbol{x}$. By moving, the point-like object generates an acceleration variation which can be expressed as

$$a_x = -\frac{m}{M}\, G_x \quad \Rightarrow \quad \delta a_x = -\frac{m}{M}\, \delta\boldsymbol{x} \cdot \boldsymbol{\nabla} G_x \sim \omega^2\, \delta x \tag{8.1}$$

The last expression is analogous if the object is a rigid body, with a finite volume, it only requires to integrate over the volume. If the displacement δx fluctuates, then a_x fluctuates. The reason why δx fluctuates, how it fluctuates, and what properties the fluctuation should have to produce the observed noise, are the points on which we work.

1. First, we observe: the further the source is from the TMs, the lesser its gravitational gradient, so it should move with a higher amplitude to account for a stronger effect. The outer parts of the spacecraft might not be impactful from this point of view.

2. The moving mass m, possibly intended as a rigid body moving altogether, can however be relatively heavy, leading to a higher force.

3. The mass m can be a rigid body moving, but the same effect can be achieved with a set of a few moving masses, or even a high number of infinitesimal moving masses. We analyze these possibilities separately.

Rigid motion. The first case is that of a rigid body moving. Let's suppose that such a body is the IBM, which is the one with the strongest gravitational stiffness $|\omega^2| \sim 5 \times 10^{-7}\,\mathrm{s}^{-2}$. To yield the measured excess noise $S^{1/2}_{\Delta g}(f) \sim 1\,\mathrm{fm\,s^{-2}\,Hz^{-1/2}}(1\,\mathrm{mHz}/f)$, the IBM motion should fluctuate as $S^{1/2}_{\delta x}(f) \sim 2\,\mathrm{nm\,Hz^{-1/2}}(1\,\mathrm{mHz}/f)$. Thus, the question translates into understanding if such a fluctuation is reasonable. For example, the most common driving mechanism of mechanical displacement is thermal expansion. Assuming a reasonable thermal expansion of the order of $10\,\mathrm{\mu m\,K^{-1}}$, the computed jitter would require thermal fluctuations of the order of $\mathrm{mK\,Hz^{-1/2}}(1\,\mathrm{mHz}/f)$. This is more than one order of magnitude higher than measurements [44], hence unlikely. This does not mean that non-thermal fluctuations might be present, but it is not straightforward to think of a mechanism generating the right spectral shape.

Other parts, located further than the IBMs, should move with even higher amplitude, even less likely. The gravitational gradient scales as the inverse cube of the distance, decaying by about two orders of magnitude across the spacecraft's length.

Thermo-elastic distortion. Another possible source of gravitational noise is that induced by thermo-elastic distortion of the spacecraft structure and equipment. In that case, the distortion would affect the entire spacecraft, possibly causing a variation of the gravitational field, larger than that of a single object moving. However, such a distortion is not straightforward to model: the shape of the deformation, the structural properties, and the profile of thermal fluctuations are not easy to model. We find that the only possible approach to this requires a realistic finite-element analysis with a realistic model of the spacecraft and thermal fluctuations. Such an analysis has been done for LPF [84], though only above 1 mHz, i.e. in the frequency band of LPF requirements. The analysis concluded that the acceleration noise induced by this effect is about one order of magnitude below the measured excess; however, no conclusions could be drawn below 1 mHz. We find that performing an analysis of this kind for LISA, extending it to lower frequencies, would be important to understand the impact on noise performances.

Step-wise motion. Since explaining a rigid motion with the right spectral shape is quite hard, a simple option could be that the movement is composed of small steps dx, singularly undetectable, occurring with a Poissonian rate λ. Before discussing the mechanical details, such a mechanism would have the advantage that its spectral shape would intrinsically be f^{-2}: as a series of Poissonian spikes generate white noise, a series of steps generate a red noise, with ASD $S_{\delta x}^{1/2}(f) \sim dx\sqrt{2\lambda}/2\pi f$, generating an acceleration noise $S_{\delta g}^{1/2}(f) = |\omega^2| S_{\delta x}^{1/2}(f)$. Even supposing a low rate $\lambda = 100\,\mathrm{d}^{-1}$, the steps would be required to have an amplitude of 0.3 nm. If so, the IBM could even move by several µm per year, getting displaced by several nm, which is unlikely, given the structural properties.

Step-wise mass loss. Another mechanism possibly leading to gravitational forces is the evaporation of particles from the spacecraft structure. This phenomenon has already been considered and acknowledged [2, 3], as it can not be entirely ruled out: for example, mass loss from structural components has already been observed in space missions; the Rosetta mission [85] estimated a loss of about 0.1 kg/yr, going on for years after launch.
This mechanism would intrinsically have the right spectral shape. Let's assume that mass is lost with small δm steps with a Poissonian rate λ, and that each mass generated a gravitational field $G_{x,i}$ at the TM location. The generated noise would be

$$S_{\Delta g} = \sum_i \frac{2G_{x,i}^2 \delta m^2 \lambda}{M^2(2\pi f)^2} \sim \frac{2\delta m^2 \lambda}{M^2(2\pi f)^2} \frac{V}{dV} \frac{\int G_x^2\, dV}{V} \tag{8.2}$$

Where $\int G_x^2 dV/V$ is the average square gravitational field $\overline{G_x^2}$. Noticing that at V/dV is the number of emitting sites N, and that $\delta m \lambda N = \dot{m}$ is the mass emission rate, the expression reduces to

$$S_{\Delta g} = \frac{2\delta m\, \dot{m}}{M^2(2\pi f)^2} \overline{G_x^2} \tag{8.3}$$

From this point, the situation can be arbitrarily complicated. We simplify it, to get an order-of-magnitude estimate. Let's assume that only one TM is present, in the center of a thin spherical shell. This is not unrealistic, as a possible emission source could be the spacecraft's outer cylinder, made of carbon-fiber-reinforced plastic. The average square TM gravitational field on a spherical shell of radius r is easily calculable, as $\overline{G^2} = \mathcal{G}^2 M^2/3r^4$, with $\mathcal{G}$ the gravitational constant. Thus,

$$S_{\Delta g} \sim \frac{2\delta m\, \dot{m}}{M^2(2\pi f)^2} \frac{\mathcal{G}^2 M^2}{3r^4} \tag{8.4}$$

This figure leads to the following conclusion: to account for the overall measured excess $S_{\Delta g}$, the emission properties should satisfy

$$\frac{\delta m\, \dot{m}}{r^4} \sim 10^{-14}\,\mathrm{kg^2\,s^{-1}\,m^{-4}} \tag{8.5}$$

This yields some conclusions:

1. If the emitting surface is close, i.e. the EH located at $r \sim 3\,\mathrm{cm}$, then $\delta m\, \dot{m} \sim 10^{-20}\,\mathrm{kg^2\,s^{-1}}$. This is unlikely: if particles were emitted with ng shots, a rate of $0.1\,\mathrm{kg/yr}$ would be necessary, which is completely unreasonable.

2. On the other hand, outer regions could emit, for instance the spacecraft's inner cylinder, made of carbon-fiber-reinforced plastic, located at about $r = 30\,\mathrm{cm}$. In that case, $\delta m\, \dot{m} \sim 10^{-16}\,\mathrm{kg^2\,s^{-1}}$. Assuming a high mass loss rate as in Rosetta, to explain LPF's noise this would require shots of $\delta m \sim 10\,\mu\mathrm{g}$, with $\lambda \sim 0.1\,\mathrm{s^{-1}}$.

We find that the last scenario is rather speculative, but it there is no reason to rule it out.

8.1.3 Magnetic noise down-conversion

Magnetic lines at high frequencies may also induce in-band force. The presence of this effect has already been raised in the past [2], but it has been investigated recently [86]. The driving reason for our interest is that magnetic effects are quadratic: specifically, a high-frequency magnetic field would induce high-frequency eddy currents within the TMs, and then exert forces on them. The effect would notably convert into force at twice the line frequency, and force at near-DC frequency, shifting amplitude fluctuations down to the sub-mHz band.
Recent calculations [86] have shown that audio-frequency lines are the ones that may possibly impact the most on the induced forces, peaking at $\sim 100\,\mathrm{Hz}$: at lower frequencies, the induction effect would be too weak, and at higher frequencies, the EHs would screen out the magnetic fluctuations. To give a scale, finite-element calculations show that a magnetic dipole of $1\,\mathrm{mA\,m^2}$, located $20\,\mathrm{cm}$ from the TMs and oscillating at $100\,\mathrm{Hz}$, would yield a force per unit mass $\Delta g \sim 4\,\mathrm{fm\,s^{-2}}$.
No high-frequency magnetometers were present onboard LPF, but tests were performed on ground before launch, showing lines with less than nT amplitudes, which could not account for the observed acceleration noise. Despite this, operating conditions might

have been different from the pre-launch ones, therefore we can not rule out this effect. Studies on magnetic effect modelization, and mitigation, are still underway within the LISA Consortium.

8.1.4 Excess noise as under-threshold glitches

In the previous chapters, we discussed the unmodeled low-frequency noise and the transient glitches (especially ICGs), concluding that both originate from true forces acting on the TMs. It is then obvious to wonder if the low-frequency noise might originate from a distribution of many glitches, under the detection threshold and not identifiable as single events. This is not a simple question in principle, as its implication would be that the noise PSD itself (hence the detection threshold) is driven by the glitch distribution properties. Our purpose here is to understand if it is reasonable that such a glitch distribution might exist, and what properties it should have to fit with both the observation of noise and glitches.

Spectral properties. First, the distribution should clearly produce a $\propto f^{-2}$ noise, down at least to 36 µHz, as the one that we observed. We only discuss glitches with the single-exponential profile, as in Eq. (7.1), as those are the most numerous, representative, and simpler to handle.
The first thing that we observe is that the detected over-threshold ICGs occurred with a Poissonian rate, implying that also the under-threshold ones could. It is known that a flow of identical spike events, occurring with Possonian rate λ and amplitude v_0, would yield a PSD $S_{\Delta g} = 2\lambda v_0^2$. Moreover, if those events were not simple spikes, but rather events with the ICG time profile whose Fourier transform modulus is $|h(f,\tau)|$, then the PSD would read $S_{\Delta g} = 2\lambda v_0^2 |h(f,\tau)|^2$. Given that the glitch Fourier transform modulus is $|h(f,\tau)| = (1+(2\pi\tau f)^2)^{-1}$ (see Eq. (7.3)), this would be the driver of the spectral shape.
From this, one may immediately draw a first, important, conclusion: only long-lasting glitches might generate a low-frequency noise. One could also conclude that a flow of identical glitches would produce a PSD scaling as f^{-4}, hence not compatible with observations. Nevertheless, that is not the LPF case, which showed instead a distribution of durations τ and amplitudes Δv.
We make two assumptions, based on LPF's observations:

- We assume that the distribution of durations was uniform in $\log\tau$, which is a fairly reasonable assumption (see Fig. 7.11), between two durations τ_1 and τ_2, hence following a distribution $p(\tau) = (1/\tau)/\log(\tau_2/\tau_1)$. Reasonable values for the durations will be discussed later.
- We also assume that the mean square impulse v_0^2 scales as τ^2, i.e., $\langle v_0^2|\tau\rangle = g_0^2\tau^2$.

With these assumptions, we calculate

$$S_{\Delta g}(f) = 2\lambda \int_{\tau_1}^{\tau_2} \frac{\langle v_0^2|\tau\rangle\ p(\tau)}{(1+(2\pi\tau f)^2)^2}\ d\tau \simeq \frac{\lambda g_0^2}{4\pi^2 \log(\tau_2/\tau_1)} \frac{1}{f^2}, \quad \text{if } \tau_2 \gg 1/f \gg \tau_1 \quad (8.6)$$

which has the observed spectral shape. Our assumptions are not the only ones possible, but are reasonable and rather simple. To account for all the measured noise

$S_{\Delta g}(f) \sim 1\,\mathrm{fm^2\,s^{-4}\,Hz^{-1}} \times (1\,\mathrm{mHz}/f)^2$, it would be required that the quantity (λg_0^2) was of the order of $10^{-4}\,\mathrm{fm^2\,s^{-4}\,Hz}$.

Parameter distributions. The last step in Eq. (8.6) assumes implicitly that $1/\tau_2 \ll f \ll 1/\tau_1$, hence requiring that glitches are long-lasting. The condition on τ_1 is easily met, as we detected glitches down to 1 s duration. Instead, the condition on τ_2, the right boundary of the logarithmic distribution, is more delicate. Indeed, to satisfy this condition down to 36 µHz, τ_2 must be at least $10^{4.5}$ s. No τs greater than $10^{3.9}$ s were detected; however, the distribution is still compatible with durations up to $10^{4.5}$ s, according to a simple Kolmogorov-Smirnov test.
Once the requirements on durations are satisfied, the question is what is the distribution of the glitch impulse $p(v_0|\tau)$ such that:

1. Its mean square value satisfies $\langle v_0^2|\tau\rangle = g_0^2\tau^2$, satisfying the condition for the production of the low-frequency noise, as the detected one.

2. It produces unresolvable, under-threshold, ICG glitches, presumably with a quite high rate so that they generate a confusion noise, quite steady in time.

3. It also produces, from time to time, identifiable and over-threshold glitches, such as the ones that have been identified on LPF, with a rate $\sim$ 1/d.

4. The over-threshold glitch distribution should not have any discontinuities in its domain, and it should be as smooth as possible.

Detectability threshold. We find that the key figure to address this problem is Fig. 7.12, representing, for each glitch, the SNR it had within its run. There, we noticed that the detected SNRs spanned from ~ 3 to $\sim 2\times10^3$, with high-SNR events gathering especially in the short-duration region $\Delta \lesssim 10^3$ s. As we said, those glitches are not important to our purposes, as they yield no low-frequency noise scaling as LPF's excess. What matters to us, is glitches with long durations. As already discussed, long-lasting glitches only occurred during the ordinary runs; in the same figure, nevertheless, we notice that there is an SNR gap at durations $\Delta \gtrsim 5\times10^3$ s, with basically no detections[1] with SNR $\lesssim 10$. It is rather unlikely that no events were measured within that SNR range just by chance.
Hence, we find that the distribution $p(v_0|\tau)$ would need to be rather discontinuous, showing events with SNR greater than 10, but at the same time showing no events with SNR within 3-10, and a rather inexplicable high rate of events below the threshold SNR = 3. Moreover, we notice again that no events lasting longer than $\tau \sim 8\times10^3$ s were detected, and those would be essential to reproduce the noise.
The distribution of the detected glitches would then be significantly separated from that of the under-threshold ones; this does not mean that the hypothesis is discarded, but it leads us to think that it is rather speculative.

[1] Only one glitch was detected with $\Delta \sim 10^4$ and SNR ~ 3, but that glitch was detected in run Mar16A, performed with `NOM` actuation settings, and very high noise conditions, which make the uncertainty of that SNR really high.

8.2 Discussion on vacuum environment

The LPF outgassing environment is quite well modeled, as we analyzed in Chapter 5. We noted that, in the long-lasting ordinary runs, the white-noise plateau level decreased as expected from pressure in a pumped system, showing that the plateau itself is indeed a good proxy for Brownian noise. Moreover, given the conversion coefficients measured on-ground, measurements allow us to indirectly measure pressure within the EH. This is very much promising for LISA, as the target pressure $\sim 1\,\mu$Pa could already be met right after about one year cruising to orbit, if the system behaved as it did on LPF.
We also showed that, unfortunately, angular measurements can not be used for pressure analyses, as the low-frequency and high-frequencies noise branches prevented the detection of the Brownian angular plateau, especially in the second half of the mission.

8.2.1 Vacuum environment and Brownian noise

In Sec. 5.2, we presented the results of the Brownian noise evolution, showing that it is compatible with an activation temperature of the order of 7 kK, and a time dependence $t^{-0.78}$.
First, the activation energy. As said, the value we found is not inconsistent with the emission of water from metals [68, 69], which is indeed the sounder explanation, as the GRSs has been exposed to atmosphere during manufacturing and assembling, and the bakeout was very mild with respect to the usual vacuum systems. The GRSs were indeed heated up to a relatively low temperature of 115 °C for a relatively low time of 24 h (also see Sec. 8.3.4), because of an incompatibility with surface photoemission and TM charge management. Usually, temperatures of 300 °C or even more are bakeout standards of ultra-high vacuum systems. It is very likely that the bakeout did not achieve a complete water extraction from the inner walls, and possibly the residual water gas within the GRS could also have been re-absorbed by the walls. Nevertheless, the activation energy value alone is not conclusive about the nature of the emission properties, more details and correspondence would be needed.

The second available parameter is the time evolution exponent. Once again, that is not conclusive about the emitting surfaces and emitted species, but it rules out many possibilities. It is also remarkable that such an exponent has been maintained for nearly two years, meaning that its source was not depleted. It is indeed common to observe different pressure decay branches with different exponents, due to different sources depleting in time.
However, we noted the detected $\gamma = 0.78$ is broadly compatible both with the emission of water from metals, and the emission of water from some LPF cables.

When describing the dependence of our outgassing model on temperature and time, we mentioned that the intrinsic dependence of temperature on time should not be neglected, in principle. We already noticed that our simplified model is effectively compatible with observations, but we also try to fit observations with a model accounting for temperature evolution.

We follow [67]. The model assumes that the adsorbing surfaces can be approximated as a mono-layer set of sites, which can both adsorb and release gas. The function $\theta(P,T)$ is the coverage, namely the fraction of occupied adsorption sites, depending in principle on pressure and temperature. When the system is pumped down with a characteristic pumping time τ, pressure is expected to respond to the following differential equation:

$$\frac{1}{P}\frac{dP}{dt} = -\frac{\frac{1}{\tau} + \frac{P_s}{P}\frac{\partial\theta}{\partial T}\frac{dT}{dt}}{1 + P_s\frac{\partial\theta}{\partial P}} \tag{8.7}$$

The term P_s is simply a parameter, related to the total number of occupied sites N_s and the chamber volume V, as $P_s = N_s k_B T/V$. We note that the equation also includes temperature variation, as variation speed dT/dt. It is easy to see that everything depends on the surface coverage $\theta(P,T)$, which is a subtle matter of discussion.
Assumptions reported in literature [68] usually employ the so-called isotherms, stating a dependence of the absorption temperature on coverage. Three choices are usually employed the most: the Langmuir isotherm, simply assuming that the absorption temperature is constant, and the rates of absorption/desorption are respectively proportional to the number of empty/filled sites; the Freundlich isotherm, assuming that the absorption temperature varies logarithmically with coverage; and the Temkin isotherm, assuming that the absorption temperature depends linearly on coverage, varying between two values T_1 and T_2. The latter is usually better representative, hence we employ that one, which leads to

$$\theta(P,T) = \frac{T}{T_2 - T_1}\log\left(\frac{1 + P/P_0\,\exp(-T_2/T)}{1 + P/P_0\,\exp(-T_1/T)}\right) \tag{8.8}$$

This equation, together with Eq. (8.7), closes the system, leaving out as unknown parameters P_s, P_0, T_1, and T_2. Reference [68] proposes, as reasonable values for water, $T_1 = 5.3\,\mathrm{kK}$ and $T_2 = 11\,\mathrm{kK}$. We integrate the differential equations, varying qualitatively the unknown parameters, to find the best-fit solution (see Fig. 8.3). We find that:

- The goodness-of-fit to long-lasting ordinary runs is not bad, though non-optimal, as the likelihood is still below 5%. It is still compatible with our previous conclusion that the white-noise plateau in cold runs was dominated by another source, not of Brownian origin.

- Parameters are quite off from expectations. Even though activation temperatures T_1 and T_2 can be chosen as consistent with expectations (respectively, 4.5 kK and 11.6 kK), our best estimate for P_0 is $2.5 \times 10^4\,\mathrm{Pa}$, which is orders of magnitude lower than the expected $8 \times 10^{10}\,\mathrm{Pa}$ (see ref. [68]). Such a low value would only be consistent with a really low and unrealistic site density, or a low and unrealistic molecular arrival time to the surface.
  The last parameter, P_s, is also not compatible with expectations: the best estimate, 6.0 kPa, with a reasonable assumption on the site density $3 \times 10^{15}\,\mathrm{cm}^{-2}$, would lead to an inner surface area of the vacuum chamber of $180\,\mathrm{m}^2$, which is quite unlikely, as it a couple of orders of magnitude higher than reality.

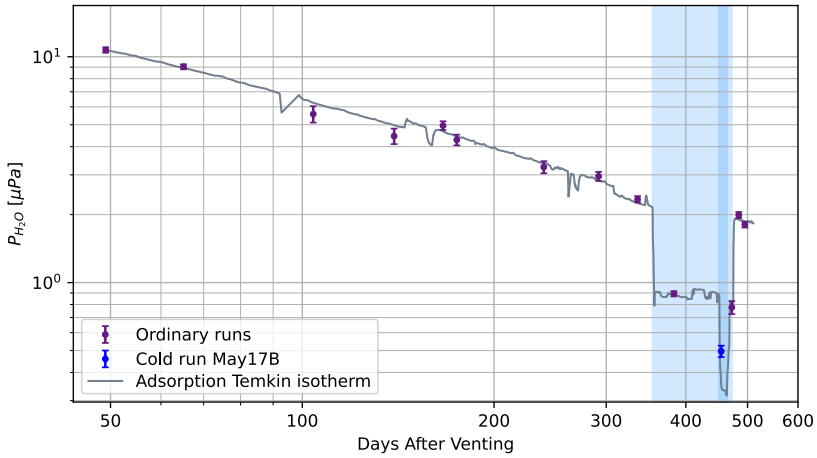


Figure 8.3: Time evolution of pressure (as in Fig. 5.1), according to Eq. (8.7), with the Temkin isotherm in Eq. (8.8), and parameters described in text.

These findings suggest that mono-layer modelization with Temkin isotherm assumption might not be the best option, but different details and assumptions might be relevant. Likely, also diffusion from the inner bulk walls can play a role, as it would intervene with a $t^{-0.5}$ time dependence; a combination of bulk diffusion and surface desorption would then occur, but that is much harder to model.

8.2.2 Long-term drift and pressure unbalance

Regarding the outgassing environment, we also found that the long-term drift detected on LPF, given its slope evolution in time, may be attributed to a consequence of outgassing. We showed that it can be simply explained as a consequence of a pressure unbalance, $\Delta P_{\mathrm{H_2O}}/P_{\mathrm{H_2O}} = (4.6 \pm 0.3) \times 10^{-3}$.
Aiming at identifying a source location providing such a pressure unbalance, we performed molecular simulations with `Molflow+` [87, 88], with a full VE geometry, rather precise in close proximity to the TM, and simplified wherever not relevant. Details about our simulations are in App. D. We observe that such a low unbalance is very hard to get if a single source in close proximity to the TM dominates the outgassing environment. For example, outgassing from IBM alone generates an unbalance higher than 10%, as the pressure on the IBM-side face would be much higher than the opposite one. This is simply justified by the fact that the impedance between the IBM side and the opposite TM face is high, because of the lack of paths joining the two regions. It is much more likely that more sources concurred jointly to the overall outgassing. Possibly, a single species dominated, as we did not observe any branching in Brownian evolution: for example, water from metal, water from polymers.
We could not perform any simulation clearly showing the observed pressure unbalance: as it is so low, such a simulation would require a very long computational time, and could also need more detail in regions that we simplified. It would also hardly be conclusive, as the parameter space of possible sources producing the unbalance is too large.

8.3 Discussion on the physical origin of LPF glitches

In Chapter 7, we discussed the properties of the transient events detected on LPF, dealing with detections, observations, and measured parameters. We have shown that LPF glitches might be divided into two well-distinct categories, and that they follow two different behaviors, during ordinary and cold runs.

We have shown that one of the two categories, named "fast, low-impulse", included rapid events, and that many of such events occurred in coincidence with differently-shaped transients in other interferometer channels. In addition to this, we also showed that these glitches occurred mainly during the cold runs, showing globally non-Poissonian occurrence statistics, and bunching up in the first stages right after re-heating. It is clear that the cold runs, performed at nominal 0 °C, showed a much-worsened behavior with respect to the ordinary ones. The effects shown there, nominally the occurrence of FLGs and the increased rate of ICGs, just add on the general behavior observed in cold runs, which were affected by increased noise in the interferometric channels, and a possibly non-thermal mechanical GRS distortion. This distortion also led to a test mass re-grab, between runs May17B and May17C. The distortion must have put under stress the interface between the GRS (basically made of metal, with a non-negligible thermal distortion coefficient) and the Zerodur^{TM} OMS Optical Bench, virtually undistorted because of the very low temperature coefficient.

We focused a lot on the "impulse-carrying" glitches, much more present throughout the entire mission, and showing well-defined templates, which would be much more impactful on the low-frequency performances in LISA. We described the templates, showed the parameter distributions, and proved that the events follow a Poissonian occurrence statistics, being hence independent of one another. Moreover, we used the LPF capacitive sensor provided by the GRS subsystem to prove that these events are real forces acting on the test masses, or, more likely on one test mass at a time. A few events had a non-zero lever arm, showing transients in the angular acceleration time series. Nevertheless, nothing relevant was found along the y and z linear degrees of freedom. No coincidences with any other LPF time series were found.

In this discussion chapter, we investigate the possible physical mechanisms that could originate the detected glitches, especially the ones related to the impulse-carrying phenomena.

8.3.1 Nature of the fast, low-impulse glitches

The nature of FLG events is much better understood than that of ICGs. Even though we were not able to firmly assess if they are true forces acting on the test masses (see Sec. 7.3.1, where we prove that ICG events are indeed true forces, but we were not able to draw conclusions about the FLGs due to strong high-frequency noise in the capacitive GRS sensor), we detected many properties that point in the same direction. We showed that FLGs did not leave any net impulse to the test masses, were high-frequency signals, and were short-lasting, hence requiring no counteraction in the feedback actuation loop. The most important feature, discussed in Sec. 7.4.1, is that the majority of such events showed coincidence with transients in other interferometer

channels.

It is highly likely that FLG events were due to anomalies in the interferometric readout, with no force applied to the test masses. We do not have particular hints about the underlying generating mechanism, but we note that those events were mainly detected during the cold runs, which notably showed an increased level of mechanical stress. Similar events (i.e., short-lasting, and transferring no impulse) also showed up in a similar interferometer flown on the GRACE Follow-on mission [89]. The interferometer readout exhibited phase jumps, that would translate for us into steps in $\Delta x_{12,\mathrm{OMS}}$. Such jumps were well analyzed during runs with an enhanced sampling rate, increased from 10 Hz to 10 kHz, and were traced back to mechanical disturbances generated by thrusters activation. Thus mechanical stress may be the root cause of these interferometer anomalies on LPF.

8.3.2 Nature of the impulse-carrying glitches

Contrarily to fast low-impulse glitches, we were able to demonstrate that a large fraction of the impulse-carrying ones consists of true force events acting on one (or both) test masses. Though we could not extend the demonstration to the smallest ones, the shape similarity with the larger ones, and the continuity of the parameter distribution, make it very likely that all glitches of this kind, or at least the vast majority of them, consist of true force events.

In the following, we deal with the possible sources causing this kind of events: first, we discuss their main statistical features, then we identify some possible sources, aiming at ruling out the unlikely ones, and identifying the most likely ones.

Impulse-carrying glitch clustering and taxonomy

We propose some considerations about the ICG observed properties, which might help in understanding the physical causes. We point out that these analyses might be found slightly more unprecise than the statistical ones presented before. Though maybe inexact, these observations are nevertheless rather objective, highlighting some points that might be relevant for a deeper understanding.

First, we have already noticed that the glitch occurrence statistics followed the Poisson statistics, during both the ordinary and the cold runs (Fig. 7.6). We observed different occurrence rates at different times in the mission, but we observed neither any significant clustering nor any repeated pattern. We have also seen glitches pile up [3] (during cold runs, when the rate was higher), and the pile-up rate is compatible with random coincidences. Thus, these glitches appear to be due to independent sources and occur at random times.

We refer again to Figure 7.7, where we observed that during ordinary runs in stable conditions, and within the specified operating conditions, we observed a constant mean occurrence rate throughout the $\simeq 1.2$ yr of the mission science operations. During this period of time, the operational settings of LPF changed many times, and pressure around the test mass decreased by almost an order of magnitude (Chapter 5, ref. [2, 5]). Apparently, the only changing element was temperature, which was lowered down to near 0 °C. We will discuss later that it was not temperature to contribute directly

to the glitch rate, but rather another mechanism, likely thermally induced. Anyway, the glitch rate increased visibly, about 30 times.

Before dealing with the cross-check with additional telemetries, we look again at the ICG glitch parameter distributions, i.e. (Δv, Δ).

More specifically, Fig. 7.11 on page 144 shows that glitches might be clustering in two roughly-different populations:

- A population with $\Delta \gtrsim 1.6\,\mathrm{ks}$, which accounts for: 37% of all positive glitches in ordinary runs; just 0.7% in cold runs; no negative glitches. Indeed, no long-lasting negative glitches were detected. The threshold 1.6 ks may be intended as subjective, we just use it as a reference for discussion, with no physical meaning.
- In both ordinary and cold runs, positive-impulse glitches constitute the vast majority of all glitches.

The number of glitches detected in the three categories is different, as shown in Table 8.1. We list the number of glitches in these three categories: positive-short, positive-long, and negative, for both ordinary and cold runs. Knowing how long cold runs and ordinary runs lasted, we make a projection of how many glitches would have been detected if cold runs followed the same statistics as ordinary. The driving question is the following: the occurrence rate of positive and short-lasting events surely increased, since the event count went from 51 to 304; however, counts of positive and long-lasting events went from 50 to 2. Is it possible that the latter distribution is still compatible with the ordinary run distribution?
We attempt a prediction: knowing that cold runs lasted 12.4 days, which is 8.99% of the duration of the ordinary runs (138 days), then we expect rates to be about 8.99% if the distribution is the same. Counts C would then be predicted according to a Poissonian distribution, as

$$p(C_{\mathrm{CR}}|C_{\mathrm{OR}}) = p(C_{\mathrm{CR}}|\lambda_{\mathrm{CR}} = \lambda_{\mathrm{OR}})\, p(\lambda_{\mathrm{OR}}|C_{\mathrm{OR}}) \tag{8.9}$$

Hence we make a Bayesian projection of the number of glitches that one would have observed if cold runs followed the same statistics as the ordinary runs, and we provide values at 90% confidence level.

Table 8.1: Observed and expected number of glitches, listed by duration and impulse sign. We recall that no long-lasting negative glitches were detected. It can be noticed that the observations for long-lasting positive glitches corresponded to projections.

	$\Delta v > 0$ $\Delta < 1.6\,\mathrm{ks}$	$\Delta v > 0$ $\Delta > 1.6\,\mathrm{ks}$	$\Delta v < 0$
Ordinary runs	51	30	17
Projected cold runs [a]	1-13	0-9	0-6
Observed cold runs	304	2	28

[a] Counts for cold runs, projected from the observed counts and rate in ordinary runs. Intervals correspond to 90% confidence.

The result of this projection shows that:

- Counts for positive glitches of duration less than 1.6 ks are more than twenty times larger than the largest projected value. Thus, cooldown has strongly increased the rate of these glitches.
- Counts for positive glitches of duration larger than 1.6 ks are compatible with the projected values, and thus the rate of these glitches has not increased upon cooldown.
- Counts for negative glitches, all of which are short, are more than four times larger than the value expected from the ordinary runs statistics. Thus cooldown has affected the rate of these glitches too, though to a lesser extent than that of positive shorter glitches.

We have shown that the increase in glitch rate during the cooldown affected basically two glitch populations, namely positive and negative short-lasting events. Another question that could arise is if the parameter distributions are still compatible, despite the rise in the occurrence rate. These considerations make no definitive proof, but they are suggestive that, in ordinary runs, ICG glitches belong to three different families:

1. Positive-impulse glitches, with $\Delta \geq 1.6\,\mathrm{ks}$.
   For this family Δv increases rapidly with Δ. For reference, we calculate that the line $\Delta v\,[\mathrm{pm/s}] = \Delta^2/\left(0.65\,\mathrm{ks}\right)^2$ is the lowest power-law upper bound to all glitches in this family. The lower bound is not precisely set, because of the presence of the LPF-noise lower detectability threshold (see Fig. 7.11). The upper bound does not appear to be due to any obvious selection effect as detection of glitches above it was not prevented for instrumental reasons.
   Remarkably, the rate of these glitches has not been affected by the thermomechanical stress conditions of the cooldown.

2. Positive-impulse glitches, with $\Delta < 1.6\,\mathrm{ks}$. This coincides basically with the glitch population of this kind during the cold runs.
   Indeed, given that the rate of the glitches of the long positive impulse family was unaffected by cooldown, even if that family extended to shorter durations, very few samples would contaminate the cold runs distribution with $\Delta < 1.6\,\mathrm{ks}$.
   About 70% of the glitches within this family last less than one minute.
   The rate of these glitches has been greatly affected by the thermomechanical conditions of cooldown. Actually, as in cold runs negative impulse glitches are only about 8% of the total, the time/temperature evolution of the rate in Fig. 7.8 refers basically to the glitches in this family.

3. Negative-impulse glitches, solely with $\Delta \leq 1.6\,\mathrm{ks}$. About 53% of these events last less than a minute.
   The rate of this family has also been affected by cooldown, though to a lesser extent. There are not enough samples to assess if the rate has been affected by temperature.

8.3.3 Possible physical sources of force glitches

We now investigate on many possible generating effects, describing the possible sources causing glitch-like events. We discuss the confidence that we have about the phenomena, based on observation and cross-telemetry checks.

Platform acceleration and inertial forces

An obvious cause of transient events, in space-borne differential accelerometers, might be a sudden acceleration of the spacecraft, which would be picked up by the differential motion because of the unperfect common mode rejection of the instrument. An obvious example of this is the case of a micrometeoroid hit.

In our case, we can rule out this source, as we corrected the data for the coupling to spacecraft motion. In addition, we have inspected the $x_{1,\mathrm{OMS}}(t)$ data series, to check for any residual coincidence, possibly due to some residual inaccuracy in the data correction, finding none. Similar data correction and inspection also rule out, as sources of glitches, inertial forces due to spacecraft rotation.

Thermal effects

We observe no coincidences between glitch occurrences and thermometer readings.

The correlation of temperature and temperature gradient variations with LPF's Δg has been investigated by a series of dedicated experiments [44, 71], aimed at measuring the dependence of Δg on the average temperature $\overline{T}$ of all thermometers on each electrode housing. Acceleration was found to follow a complex behavior, with a relatively prompt response and coefficient $\partial\Delta g/\partial\overline{T}|_p$, and an additional delayed response with a different coefficient $\partial\Delta g/\partial\overline{T}|_d$, likely due to the delayed heating of distant sources. The former was found to be pretty constant over time, at about $0.5\,\mathrm{pm\,s^{-2}\,K^{-1}}$. The coefficient $\partial\Delta g/\partial\overline{T}|_d$ was instead found to decrease by about a factor 5 over the course of the mission and to level off at $\simeq 0.5\,\mathrm{pm\,s^{-2}\,K^{-1}}$, paralleling the decrease in pressure.

A similar pressure-dependent behavior was also found for $\partial\Delta g/\partial\Delta T_{\mathrm{EH}_i}$, where ΔT_{EH_i} is the difference of temperature across the ith electrode housing. For both electrode housings, these coefficients leveled off at $\simeq 10\,\mathrm{pm\,s^{-2}\,K^{-1}}$.

With such sensitivities, to explain the smallest of the observed Δg glitches with a glitch in $\overline{T}$, one would need amplitudes of order mK at the beginning of the mission and of many tens of them at the end. We would have detected such glitches, as our resolution is of the order of tens µK in $\overline{T}$ for a 100 s glitch following the template in Eq. (7.1).

The same applies to a glitch in the differential temperature ΔT_{EH}, for which glitch amplitudes would need again to be of order of mK, and where our sensitivity is in the µK range thanks to a dedicated low noise temperature differential readout [44].

Therefore these hypothetical temperature glitches would have been detected within the temperature time series. We believe that this rules out the hypothesis that glitches may be due to thermal transients.

Gravitational signals

The possibility that glitches originated from some amount of mass permanently changing position or leaving the system is easily ruled out, as they leave no residual acceleration after their occurrences. Hence, we have to seek for a different origin.

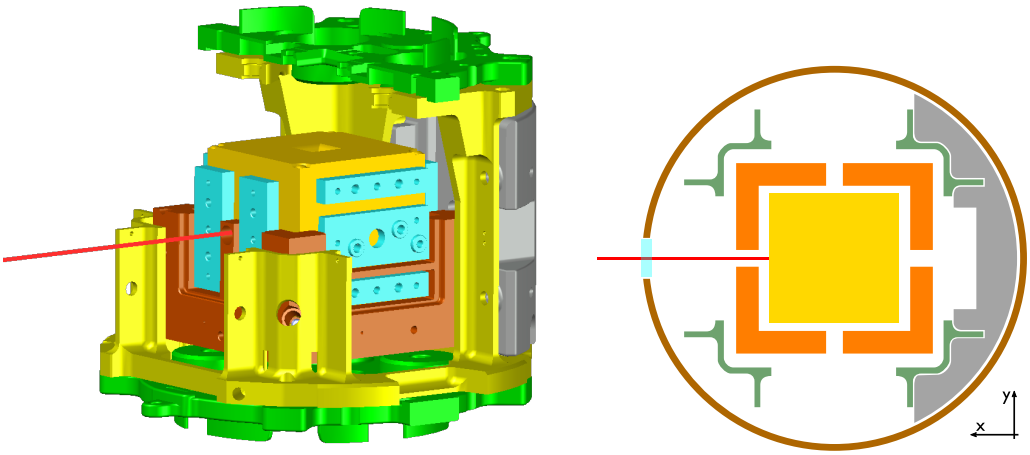


Figure 8.4: Schematic representation of the GRS, 3D visualization (left) and horizontal section (right). The golden cube is the Au-Pt test mass. The orange/brown hollow structure represents the electrode housing, whose mid-section carries four symmetric holes, one of which is the input port for the laser beam (red). The 3D view also represents electrodes. The grey element is the specially shaped tungsten gravitational balance mass.

Another possible cause for a gravitational signal would be a body moving around its equilibrium position, moving exactly with the time profile of Eqs. (7.1, 7.2, 7.5), settling back to the initial position after its transient motion. An example could be the 2 kg tungsten balance mass located within each GRS, at a few centimeters from the test mass center (see Fig. 8.4).

The gravitational force gradient per unit mass acting on the test mass due to the balance mass is $-\omega_{1,2}^2 \simeq 5 \times 10^{-7}\,\mathrm{s}^{-2}$. Thus, to produce our smallest glitches with a peak amplitude about $\mathrm{fm\,s}^{-2}$, one would need a displacement of the balance mass out of its equilibrium position peaking at $\sim 2\,\mathrm{nm}$ and then getting back to rest.

The balance mass is indeed the largest source of gravitational gradient. The gradient from other sources is significantly smaller, as the gradient from the rest of the GRS drops to about $4 \times 10^{-7}\,\mathrm{s}^{-2}$, due to compensation [50], and as the contribution of farther apart components decays as the cube of the distance. For instance, the optical bench contributes a gradient of about $6 \times 10^{-8}\,\mathrm{s}^{-2}$, the other GRS $3 \times 10^{-8}\,\mathrm{s}^{-2}$, and the rest of the spacecraft $1 \times 10^{-8}\,\mathrm{s}^{-2}$ [90]. Those parts would be required a larger displacement, to yield the same gravitational effect.

The question, then, is what could cause such motion of mechanical part, and how it could explain the long-lasting profile. Indeed, the characteristic frequencies of the mechanical parts surrounding the test mass are in the kHz range: even the overdamped mechanical motion of those parts takes place on timescales much shorter than that of the vast majority of the observed glitches. In order to produce a glitch by way of this physical mechanism, one would need parts to actively move because of

some transient distortion, and not by some mechanically excited free motion.

We foresee two major patterns for such a thermomechanical distortion. The first is an expansion or contraction of the GRS around the test mass due to a temperature fluctuation. The major effect of such distortion would be to move the balance mass relative to the test mass. Given the construction details, this effect adds a term of order $\simeq 0.3\,\mathrm{pm\,s^{-2}\,K^{-1}}$ to the temperature coefficient $\partial\Delta g/\partial\overline{T}|_p$ discussed above. The possibility of glitches originating this way has been already ruled out.

The second distortion pattern that may give origin to a force transient, is a rigid displacement along x of either of the two GRS relative to its own test mass. This is the case, for instance, if the struts that attach the GRS to the LTP bay thermally expand or contract, moving both GRS in opposite directions, and changing their distance by some amount δl.

By moving the main sources of gradient, such a distortion would cause both a signal $\Delta g = -\omega_{2,\mathrm{GRS}}^2\,\delta l$, with $-\omega_{2,\mathrm{GRS}}^2 \sim -\omega_2^2 \sim 4\times10^{-7}\,\mathrm{s^{-2}}$, and a signal $x_{\mathrm{OMS}} - x_{\mathrm{GRS}} = \delta l$. As already mentioned, we have observed both such signals upon the large distortion caused by the cooldown.

Thus, a glitch originating from such a distortion pattern, or from any pattern that would move any of the GRS relative to its test mass, should also show up in the $x_{\mathrm{OMS}} - x_{\mathrm{GRS}}$ series, peaking at $\Delta g_{\max}/(-\omega_{2,\mathrm{GRS}}^2)$, a number that ranges from a few nanometers to micrometers. We have inspected the series $x_{\mathrm{OMS}} - x_{\mathrm{GRS}}$ and found no corresponding glitch.

A distortion pattern moving the spacecraft, or one of its large components, relative to the entire LTP and nondetectable in $x_{\mathrm{OMS}} - x_{\mathrm{GRS}}$ would require much larger amplitudes. The difference of gradient on the two test masses due to the entire spacecraft, which is the quantity that would matter in this case, is of order $10^{-9}\,\mathrm{s^{-2}}$. Thus, to produce the observed glitches, the spacecraft would have to have moved along x by an amount ranging from micrometers to millimeters, requiring temperature changes ranging from a fraction of a K to hundreds of K, without causing any detectable distortion within the LTP. We consider such a scenario to be quite unlikely.

In conclusion, we believe that a gravitational origin is an unlikely explanation for the vast majority of the observed glitches.

Magnetic force

It also appears unlikely that these force glitches are explained by some slow transient in the magnetic field. With *slow* here we mean that we are not considering eddy current effects, which we will discuss later.

The magnetic susceptibility of the test masses has been measured to be $\chi \sim -2.5\times10^{-5}$ and its permanent magnetic moment $|\boldsymbol{\mu}| < 5\,\mathrm{nAm^2}$ [15], though this is just an upper limit. The static magnetic field on board LPF was found to be $|\boldsymbol{B}| \sim 1\,\mu\mathrm{T}$ [72], and, finally, from the lack of correlation between the magnetic field and force noise [5], we estimate the magnetic gradient to be less than $10\,\mu\mathrm{T/m}$.

The force peak amplitudes of all observed glitches have $\Delta g_{\max} \geq 10^{-15}\,\mathrm{m\,s^{-2}}$ (see Fig. 7.14). To reach such a force level, a glitch in the magnetic field gradient, on either of the test masses, considering only the coupling to the test mass induced

moment, should peak at about 1 µT/m. Unless we take into account an unrealistically close source, a gradient like this would have produced a detectable signal on some of the magnetometers, at least for glitch duration larger than about 10 s. As glitch amplitudes were even 1000 times larger than this minimal one, many of the glitches would have produced large magnetometer signals.

Similar conclusions are obtained assuming that the glitch is in the magnetic field. To produce our force glitch with a glitch in the magnetic field, one would require a peak at least 0.1 µT, well above the detection threshold. Thus the lack of observed magnetic glitches rules out this explanation.

A further mechanism for force transient of magnetic origin, is that of currents induced via Seebeck effect by thermal gradients inside the test mass. The test mass material has a finite amount of Au and Pt precipitate that may create effective thermocouples at grain boundaries. The effect was noticed during magnetic characterization of the test masses on ground, during which temperature differences of order of a few K across the test mass, due to manipulation by human hands, induced a magnetic moment peaking at some $\mu A\,m^2$. With the value for the static gradient quoted above, this effect would just give a correction of $\simeq 0.1\,pm\,s^{-2}\,K^{-1}$ to the coefficient $\partial\Delta g/\partial\Delta T_{EH_i}$ mentioned in the previous section, and its role in producing glitches is ruled out together with the rest of the thermal effects.

In addition to these low frequency effects, magnetic fields at high frequency may induce eddy currents within the test masses and then exert Lorentz forces on them [14]. The effect is thus quadratic and would convert the low frequency amplitude fluctuations of a high frequency magnetic spectral line into a corresponding low frequency force.

To give a scale of the effect, a recent finite-element electromagnetic calculation [86], has shown that the effect of a dipole of $1\,mA\,m^2$ located at a distance $d = 20\,cm$ from the test mass and oscillating at the frequency of 100 Hz, would cause a force of $\Delta g \simeq 4\,fm/s^2$. The effect reaches its peak at 100 Hz, while at lower frequencies the induced current decreases and above that the screening effect of the metallic electrode housing attenuates the oscillating field.

The effect of a dipole source decreases with d^{-7}, so that at the closest distances of about 0.4 m between the test mass and any active device on the LPF spacecraft the effect might be ~ 100 times smaller.

However, the shape and timescale of the glitches are not easily reconciled with such a magnetic origin. Electromagnetic emission from electronics is usually modulated by noise, by switching among different operational settings, and by thermal variations, and we do not see how these may easily follow the reversible exponential behavior lasting minutes to hours that would be required to generate the observed glitches.

One way though of producing a smooth time evolution is that of two lines of constant amplitude, the frequencies of which would slowly drift over time. If during some time interval these lines had a substantial overlap in the frequency domain, they would indeed generate a force on the test mass.

Lines observed during testing where stable in frequency, but, again, we cannot exclude that other lines were present in flight. It seems, however, highly unrealistic that on orbit enough lines have been generated, with different enough drift rates and shapes, to explain the hundreds of glitches of Fig. 7.11, with parameter values that

span a few orders of magnitudes.

While the possibility that some of the observed glitches are due to eddy currents cannot be discarded, it is highly unlikely that this source may explain the majority of the observed glitches.

Electrostatic forces from GRS electronics

Another option is obviously related to electronics, and electrical forces in general. The simplest explanation would be for example a cosmic ray event depositing a high charge on the isolated TMs. Though, it is quite easy to exclude such an occurrence, as a deposited charge would leave a permanent acceleration step in Δg, quite incompatible with the observed finite-impulse glitches. Moreover, the rate would not change with temperature, but rather change with solar wind intensity, which is not the case. In addition, all surfaces facing the test masses are conducting and grounded, and would not accumulate free charge.

Still, glitches may be produced by spurious voltage transients in the electronics used to control the test masses. More specifically, the mentioned electrodes facing the x-faces of the test masses, are all driven by separate amplifiers. A voltage glitch in one of these amplifiers would certainly produce a force on the test mass.

However, a single-electrode event like this would also generate a torque around z, with a lever arm $|r_\phi| = 11\,\mathrm{mm}$, which is the main reason why we have searched for a torque component to the detected force glitches as explained in Sec. 7.3.2.

To span the observed range of glitch peak amplitudes, from $1\,\mathrm{fm/s^2}$ to $1\,\mathrm{pm/s^2}$, with the electrode geometry, there would need to be a transient change in the mean square voltage at the actuation amplifier output between roughly $10\,\mathrm{mV^2}$ and $10\,000\,\mathrm{mV^2}$. This could occur in different forms:

- A transient "quasi-DC" voltage in the 100 µV to 100 mV range, mixing with stray DC potential differences of order 100 mV, due to test mass charge and/or stray "patch" voltages.
- A transient change in the roughly 1 V actuation audio-frequency carriers [1] by roughly 10 µV to 10 mV.
- A transient electrode oscillation coherent with the 100 kHz sensing "injection" frequency [16] and mixing with the 0.6 V amplitude test mass bias, in roughly this same amplitude range.
- A spontaneous AC oscillation, at some random frequency not associated with the actuation or injection, with amplitude in the 5 mV to 200 mV range.

While we cannot directly exclude any of these – though the 100 kHz excitation would have likely given some capacitive sensing error – they were not detected in dedicated preflight tests, albeit relatively short (less than day per electrode), which could have detected such anomalies.

A strong indicator that the glitches do not originate in the actuation electronics comes from the analysis of the possible torque component to the observed glitches. The findings in Sec. 7.3.2 show that:

- There are in total 56 glitches within ordinary runs, spread over the entire parameter space, for which we would have been able to detect a lever arm of 11 mm (see Fig. 7.18). Only one of these is both incompatible with $r_\phi = 0$ and compatible with $r_\phi = 11\,\mathrm{mm}$ (see Table 7.2). Though there is no proof that this glitch is indeed of electrical origin, one might nevertheless take 1/56 as a rough bound to the fraction of glitches that may be due to this source.

- For the cold runs glitch data, none of the 147 glitches for which there is sufficient resolution to resolve $r_\phi = 11\,\mathrm{mm}$ have such an effective arm. The probability of such an event, using binomial statistics, and assuming the distribution is the same as during ordinary runs, is $p = 0.08$, a figure that does not allow us to reject the equal distribution hypothesis. Using both observations, ordinary and cold runs, the probability of such an occurrence becomes $p \leq 0.023$ with 95% confidence.

- For the kind of standard, audio-frequency electronics we are discussing here, minute to hour-long transients which are not induced by some corresponding thermal transients, are quite unexpected. Of the 121 glitches with $\Delta \leq 1\,\mathrm{min}$ and detectable 11 mm lever arm, none is found to have such a lever arm, which gives $p \leq 0.024$ with 95% confidence.

While this effective arm test is inconclusive for the smallest and fastest glitches, for which our sensitivity to a lever arm is reduced, most of our glitches are incompatible with a single-electrode electrical origin.

In addition to this, even for the smaller, faster glitches, the observed increase in rate upon cooling the spacecraft is not easily reconciled with an electrical origin.

More complex voltage events, simultaneously affecting more than one electrode – such as two adjacent electrodes which combine to give force without torque – are even less likely given the design of the electronics [47].

8.3.4 Outgassing environment, glitches as gas bursts

One candidate source of force is the exchange of momentum between the test mass and the gas molecules surrounding it, the same possibility that we also analyzed from the point of view of low-frequency noise. We analyzed carefully this option, which right now is the only candidate not excluded yet.

The driving observation is that the vacuum environment surrounding the test mass is rather unusual, as the vacuum chamber is densely packed with components: test mass, electrode housing, test mass launch-lock mechanism, various cable bundles, etc., as can be noticed in the left panel of Fig. 8.4. Thus, the outgassing surface-to-volume ratio is unusually high for a vacuum system, and the distribution of outgassing surfaces is rather non-symmetric around the test mass.

In such an environment, one possible source of glitches may be an event of release of some metastably trapped gas from pores. Similar events are often observed in vacuum systems, due to so-called virtual leaks – cavities with a high impedance connection to the outside – that may trap gas and release it in bursts [91]. The phenomenon is also known to be triggered by mechanical stress and friction. Pressure

events, roughly similar in shape to LPF's and lasting between tens and hundreds of seconds have already been detected in vacuum systems, as can be seen in reference [92], though in different experimental conditions, and with higher amplitudes than LPF's. Nevertheless, in that reference, glitch events were discarded and not further examined.

Transients during bakeout. We re-analyzed the pressure gauge time series during vacuum preparation of the GRS on the ground (see Fig. 8.5), which indeed showed transient long-lasting events. For that test, the GRS had been inserted in a wider vacuum chamber with its venting valve open. The chamber was pumped down and its temperature was kept at $\simeq 115\,^{\circ}\mathrm{C}$ for about 24 h. This is a standard procedure known as bakeout, meant to get rid of most of the adsorbed water.
The time stretch containing the glitches in Fig. 8.5 was observed during the final cooldown, during which the system was subject to a significant amount of thermomechanical stress due to the relatively rapid contraction. Similar behavior with many spikes was also observed during the preparation of the other GRS, while no spikes were detected during a test with the empty chamber. However, we have no way of assigning with certainty the source of these gas emission spikes to the GRS interior, and we only show them here as an example of the phenomenon in vacuum systems.

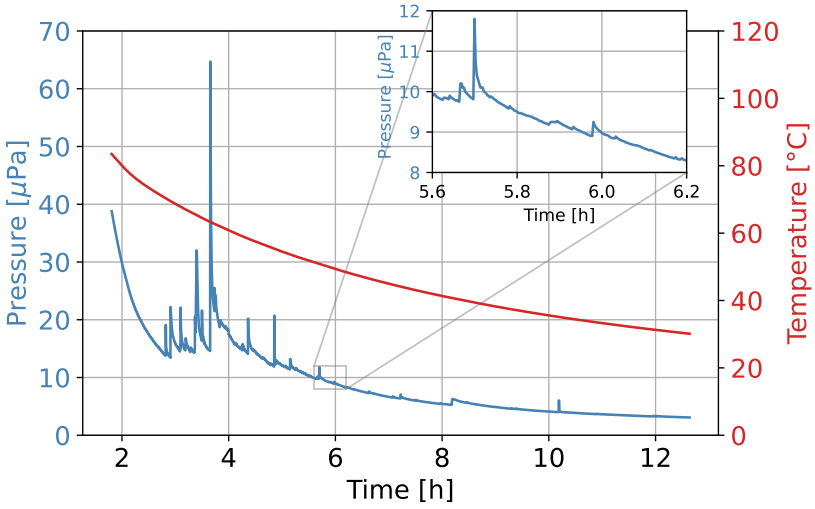


Figure 8.5: Blue line and left scale: pressure in one of the two GRS during cooldown, after about 24 h pumping at about 115 °C for vacuum preparation (bakeout) on ground. Red line and right scale: temperature in the test facility. Vacuum preparations were performed by inserting the GRS inside a wider vacuum chamber with its venting valve open. The pressure and temperature shown in the figure are those of this wider chamber. The time origin is set at the end of the bakeout phase. The inset zooms on a region with low-amplitude events. (Data courtesy OHB Italia)

Molecular simulations for ICG events. We have performed molecular dynamics simulations, with the tools provided by `Molflow+`, on a more realistic model of the GRS – See App. D for a description of the software.
Molecular simulations on simplified geometries had already been performed in the past [3, 22, 93], showing that a molecule released from a surface nearby the test mass

can transfer a net momentum to it before leaving the system via the venting duct. For instance, a water molecule with a Maxwell-distributed momentum, entering the laser port in the center of the x-face of the electrode housing (see Fig. 8.4) and hitting the test mass, would exchange with the test mass an average momentum per unit mass along x of $\Delta v \simeq -2.4 \times 10^{-22}$ m/s, before leaving the electrode housing through any of its holes. This is about a factor 40 larger than the mean momentum of the distribution $\langle \Delta v \rangle \simeq 5.7 \times 10^{-24}$ m/s, an enhancement due to multiple collisions between molecule and test mass, caused by the constrained geometry [22], and the re-thermalization at each hit.

Using the figure above for the impulse per molecule: the glitch with the largest test mass impulse $\Delta v \simeq 1$ nm/s would correspond to roughly 4×10^{12} water molecules (0.13 ng) hitting the test mass; the sum of all glitches in ordinary runs and during the entire mission, assuming the observed constant rate, would amount to about 1.5 ng; those observed during cold runs would amount to about 0.15 ng. These figures constitute a comparatively small amount of molecules, when compared with the total outgassing rate from each GRS, of the order of 100 µg/d at the beginning of the mission and 10 µg/d at the end.

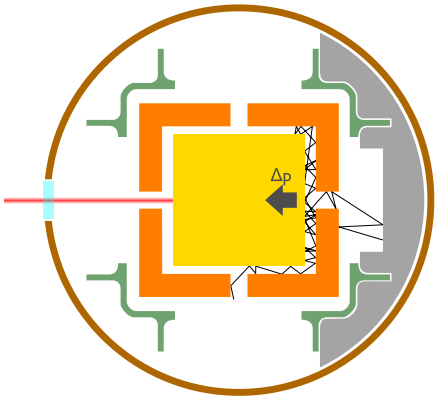


Figure 8.6: Horizontal section of the GRS, simplified, showing a possible path of molecules emitted by the IBM, hitting prevalently the $x+$ face.

We have simulated with the full GRS geometry, an instantaneous emission of molecules from a point on the side of the tungsten mass that faces the aperture in the center of the x-face of the electrode housing, identical and opposite to the mentioned laser port (see Figs. 8.4, 8.6). This emission indeed creates a glitch in the differential pressure with a shape similar to the observed ones. In particular, the profile never crosses the line $\Delta g = 0$, meaning that the TM side facing the emission region is preferentially hit.

The timescale and the profile shape are not straightforward to guess, as they heavily depend on the system geometry and response. In particular, the characteristic permanence time of molecules in such an ultra-high-vacuum system is less than 0.1 s. Hence, any long-lasting pattern should be given by either *a)* a long-lasting emission pattern, or *b)* a long mean sojourn time of the molecules on the various surfaces.

In any case, the total transferred momentum Δv is independent of the event timescale. With $m_{\rm m}$ the total mass of the emitted molecules, we find $\Delta v / m_{\rm m} \sim 0.5\,{\rm nm\,s^{-1}\,ng^{-1}}$. Note that the simulation shows that only 20% of the molecules emitted from the balance mass indeed enters the electrode housing, while the others follow different paths.

In Fig. 8.7 we provide the results of a molecular emission from the front face of the IBM, showing the different pressure profile arising on the $x+$ and the $x-$ TM sides. The simulation is run with a sojourn time $\tau = 1$ s, which is necessary to obtain a long-lasting glitch. Otherwise, with no sojourn time, the glitch would last only a

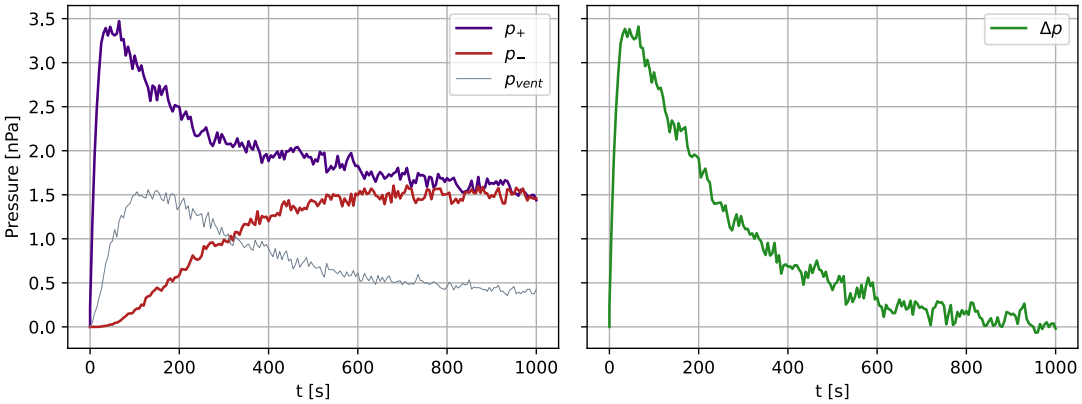


Figure 8.7: Molecular emission from the front face of the IBM, pressure profiles on the $x+$ and x- TM sides (blue, red), difference (green) and pressure at venting duct (grey). Curves are re-scaled so that the total transferred impulse is 1 nm/s. The simulation is run with a sojourn time $\tau = 1$ s, which is one of the two conditions to obtain a long-lasting glitch. The other possibility, which we find sounder, is that the emission timescale is the driver of the glitch timescale.

few tenths of a second. In the plot, curves are re-scaled so that the total transferred impulse is 1 nm/s. The grey line represents the pressure increase at the venting duct entrance, which in this case would be of the order of a few nPa.

Molecular simulations for anomalous events. What we were not able to reproduce with simulations are the few detected oscillating glitches, such as the one in Fig. 7.5. Such an event could be reproduced in the following way: the gas source emits an outgassing flow $\dot{Q}(t)$ shaped as $h(t)$ in Eq. (7.1); the impulse response function for that location is, for geometrical reasons, a linear combination of $\delta(t)$ and $\delta'(t)$, so that the observed pressure is

$$p(t) \propto h(t) * \left[\delta(t) + \beta\delta'(t)\right] \sim h(t) - \beta\dot{h}(t) \tag{8.10}$$

Such a linear combination could arise, for example, from the existence paths leading to the opposite TM side before the front-facing one. The previous expression, indeed, would coincide with observations Eq. (7.6). Unfortunately, that is not the case for LPF. Probably because of the presence of the mechanical structure, yellow in Fig. 8.4 on the left, and the absence of gaps within it and the VE, with simulations we find no locations producing "opposite patterns": if the source is located on the $x+$ (or -) TM side, it produces an effective pressure on the + (or -, respectively) side. If the source is located on the y or z sides, the glitch is prevalently produced in the y and z directions, and would likely be detected with capacitive measurements.

Lever arm of gas outflow. Note also that simulations show that the lever arm of a gas inflow from the main inlets to the electrode housing, like that coming from the balance mass, is negligible. Thus, for the few observed glitches with nonzero lever arm, one should assume that the gas has been emitted by some source localized inside the electrode housing.

However, no other source creates the same sort of cavity around one of the electrode

housing apertures as that created by the special shape of the tungsten balance mass. This is reflected in the fact that, for all other sources, the ratio between the number of molecules hitting the test mass and that of those following different paths is always significantly smaller than that for the tungsten balance mass.

Outgassing candidates. The tungsten balance masses are a natural candidate for such gas-burst events. First, a microscope analysis has shown that its sintered material is porous, with micrometer size pores [3]. Second, the sign of the transferred impulse for this source would be positive, as for the great majority of the observed glitches.
Many other components may also trap gas, beginning with the various bundles of cables that connect the electrodes, and the various motors of the launch lock, to their respective electronics. Some of these sources have the proper position to create also negative-impulse glitches that indeed we are able to reproduce with Molflow.

One more argument that may support the gas-release interpretation of glitches is the sensitivity of gas emission to thermomechanical stress, as is well illustrated by Fig. 8.5. As said, thermomechanical stress accompanied the cold runs and their highly increased glitch rate. In particular, Fig. 7.8 shows that the occurrence rate switched almost reversibly, from the $\lambda \sim 1\,\mathrm{d}^{-1}$ of ordinary runs, to the many tens per day of the cold runs, when crossing a comparatively narrow temperature range of a few degrees. In addition the data are also suggestive of some slow transient relaxation at the lowest temperatures. This may indicate that the rate may be following some complex, non-linear stress pattern due to the significant differential thermal contraction of the high numbers of equally complex contact interfaces within the GRS.

Glitch duration in case of gas bursts. Though the likelihood of this source looks the highest among those discussed so far, a few aspects remain to be clarified. An important one is the time profile of the glitches and the associated wide distribution of Δ. Most of the spikes in the data of Fig. 8.5 have an almost instantaneous onset, followed by the decay pattern one would expect in a standard vacuum system: large gas releases saturate the adsorption speed of the chamber solid surfaces and are quickly pumped down, while smaller ones follow a slower decay in quasi-equilibrium with surface adsorption. This is the response one would expect for a fast, virtually instantaneous, release of gas from some pocket.

In the case of an instantaneous emission of a group of molecules from a specific source, TM acceleration would also consist of a rise followed by a decay. The rise time would be due to the distribution of diffusion times from the source to the electrode housing inlet, and the decay time to the diffusion of molecules. The time constant of such decay would nevertheless be fixed, depending on the mean sojourn time of particles on walls and on the number of wall hits. Sojourn times of molecules on metal surfaces depend exponentially on their binding energy (see App. D). Thus they may vary by orders of magnitude, from $10^{-12}\,\mathrm{s}$ to more than seconds, depending on the nature and state of the surface, the nature and amount of adsorbed species etc. However, once the sojourn time for the electrode housing is fixed, the resulting duration for such a particular match is fixed. Indeed, the duration of the decay branch depends only on that choice, and for $\sim 90\%$ of the observed glitches the template

is that in Eq. (7.1), which only contains one time constant, so that the duration of the decay branch fixes the overall duration. The model of an instantaneous release of gas from one source would then not reproduce the observed large variability of glitch duration.
As a consequence, if glitches are due to gas release, for a large fraction of them their time profile must be dictated by the intrinsic time evolution of the gas release, while it can be limited by diffusion across the GRS only for some of the shortest ones.

We note that standard diffusion in simple geometries, like from the bottom of a very narrow pit or from the center of a spherical piece of material, does produce a time evolution of the gas outflow very close to that of our observed glitches, and that the timescale for diffusion may indeed be very long, depending on the gas species and the material.
However, we were not able to find in the literature any reference to events of slow gas release from pores or other imperfections. The only hint we have of the possibility of some non-instantaneous gas evolution is in very few of the peaks in Fig. 8.5, showing indeed some minute-long rise times.

Glitch amplitude in case of gas bursts. Moreover, the scale of the events in Fig. 8.5 is orders of magnitude different in amplitude from the kind of release that would explain our glitches. For instance, the above-mentioned Molflow simulation shows that the 2 ng of total water molecules emitted from the balance mass, needed to generate the glitch with the largest Δv, would generate a peak pressure of a few nPa at the venting valve, well below the measurement resolution of the figure.

Lastly, we find no sound explanation for the near-quadratic dependence of the impulse on the duration of the event. From the perspective of gas-bursts, the impulse would be related to the total number of molecules hitting the test mass.

8.3.5 Impact of LPF glitches on LISA

Assessing the impact of LPF-like glitches on LISA science is not straightforward, and it is a task that is currently being carried out within the LISA consortium [76, 94]. That analysis is not our purpose here, though we have a look at the main points.
As we said, the major impact on science would be made by ICG-like glitches rather than FLG-like, as the former would affect the low-frequency branch. Fast glitches, like FLG which carried no impulse, would be quite easily identified and distinguished from the GW waveforms of inspiralling astrophysical sources. On the other hand, ICG-like glitches could affect the science $\sim$ mHz frequencies, potentially spoiling observations. Analyses are being carried out within the context of LISA simulations and the LISA data challenge, injecting LPF-like glitches in simulated data and looking at how their presence modifies GW parameters. The longer the proper duration of the glitch, the higher spectral power it would have at low frequencies. As one could guess, the key point is not necessarily the duration of the glitch, but rather *when* the glitch occurs during the evolution of a binary system. As the LISA sources evolve slowly in the low-frequency band, staying even for weeks or months in it, the glitch could happen in a relatively quiet moment of the evolution, as well as it could happen in the worst moment, which is the merger.

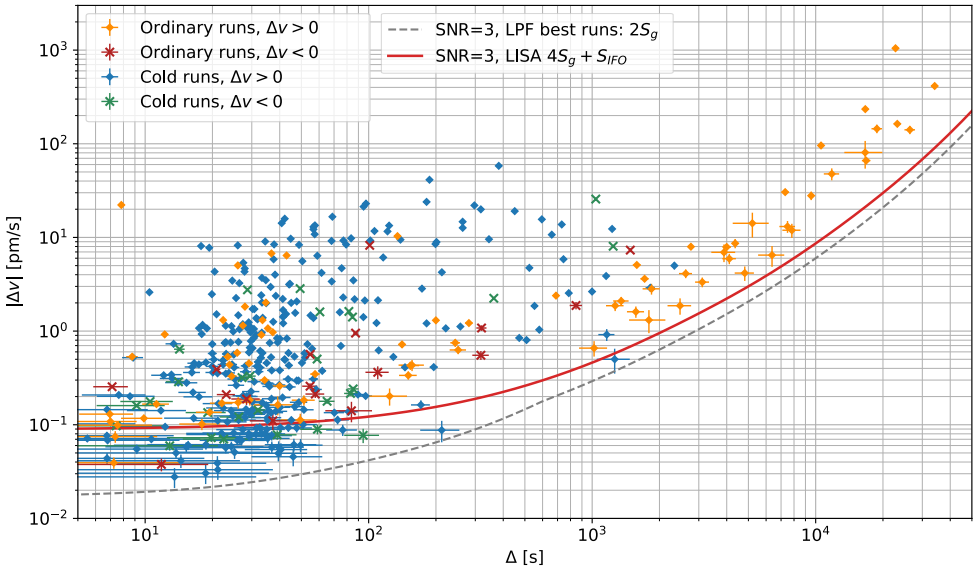


Figure 8.8: Collective plot of all LPF's ICG glitches, points, and LPF's at SNR = 3 threshold line during the best-noise runs. The red line is a rough estimate of the two-links LISA detection threshold, based on the requirements in Eqs. (1.3) and (1.4). The difference at low durations, below $\Delta \sim 10^3\,\mathrm{s}$, is driven by the higher interferometer noise, increasing from LPF's short arm $32\,\mathrm{fm\,s^{-2}\,Hz^{-1/2}}$ to LISA's long arm $10\,\mathrm{pm\,s^{-2}\,Hz^{-1/2}}$. We recall that this estimate does not take into account many variables (TDI, LISA dynamics, etc.), that should be considered in a precise analysis.

We highlight a few points about the glitch mitigation task:

1. LISA data analysis is completely different from LVK's. Glitches in on-ground interferometers occur frequently, superimposing to the background noise. This allows to better characterize them, especially when they have a high SNR. Most of the time, measurements are simply (stationary) noise, with no GW sources. Hence, the probability of having glitches superimposing to GW waveform is really an unlucky case. Though it happened for one of the most important observations, GW170817 [95], during which one of LIGO's detectors was affected by a large glitch. In that case, though, another independent detector was online, allowing a precise estimate of GW parameters. LISA will not have multiple independent detectors.

2. LISA will not detect glitches in the same way as LPF did. Analogously to LVK's observation, LPF's was a pure measurement of the out-of-loop differential acceleration between the TMs. LPF detected glitches showing up over its stationary noise, allowing us to precisely characterize them. LISA will perform interferometric-precision measurements only between distant TMs, affected by long-arm issues. Single, one-arm measurements do not allow any significant measurements, as they are heavily affected by laser frequency noise, requiring a further post-processing step known as Time Delay Interferometry (TDI). We do not go into the details here, but the relevant point is that this procedure requires the combination of many-link measurements. The post-processing, then,

might even hinder the information about the TM that produced the glitch, and will also change dramatically the glitch shape.

3. This does not mean that single-spacecraft measurements are not feasible, but rather that they are not feasible at a high-precision interferometric level. Indeed, single-spacecraft acceleration could only be detected with GRS capacitive readings [96], much noisier than the interferometric ones, hence only useful for high-SNR events.

4. Arguably, glitches occurring close to the merger of two astrophysical objects would majorly impact the GW parameter estimation. On the opposite side, glitches superimposing to, for instance, the long single-frequency inspiral of a galactic binary, would not affect greatly the measurement. Analyses are still underway in that sense [76, 94].

Even though, as said, precise analyses are much more complicated than this, we can make a very rough estimate, approximating the SNR that LPF-like ICG glitches would have in the LISA case. The need for TDI means that at least four laser links are necessary, hence meaning using four TMs. Figure 8.8 shows a collective plot of all LPF's ICG glitches, with a forecast based on the LISA requirements Eqs. (1.3) and (1.4). Note that this is just a rough estimate as, in the real case, making TDI combinations would require at least four links, hence at least double the noise.

8.4 Conclusions

Within this thesis work, we encompassed a number of aspects of the LISA Pathfinder mission, trying to extract all the information and science from data. What matters the most is that LPF succeeded in showing a residual acceleration better than the requirements for gravitational wave observation, clearing the way to adoption.
Still, most of the science that was enabled by LPF is not attainable on-ground, and as such will not be accessible before LISA will get operational. Hence, the need to carefully analyze data from LPF, and to get as much as possible from LPF, as a large part of the LISA instrumentation will directly inherit from LPF.
In-depth analyses of LPF measurements allowed us to extract some more details about its behavior, and its performance. Due to its high sensitivity, LPF uncovered some aspects that still remain not fully explained, which will also guide future on-ground analyses before LISA's launch.

As all over the place in this work, we also divide into three parts this concluding section.

1. LPF showed a low-frequency sub-mHz acceleration noise, below GW requirements but exceeding modeled predictions.
   First, we analyzed it to understand its nature, unconcerned about its physical source. We found that noise has been *a)* stable to within 25% fluctuations all over the mission, without showing any evident noise epochs across more than one year *b)* Gaussian, over the considered frequency set, *c)* caused by real forces acting on the TMs, not by an error in the interferometric readout. We also studied the angular acceleration time series, showing that the lever arm of the residual acceleration is not compatible with what would result from single-electrode disturbances.
   We analyzed the correlation between the detected excess with a number of diagnostic telemetries and dynamical quantities, known to exert spurious forces on the TMs, namely temperatures, thermal gradients, magnetic fields, actuation digitization residuals, fuel mass loss, instrument distortion. We found that, in the band [36 µHz–1 mHz], none of these sources dominates observations, accounting overall for less than 10%. On the other hand, the frequency 18 µHz is dominated by thermal fluctuations, especially at the beginning of the mission.
   We gave a projection of the unaccounted effects contributing to LPF noise. At the present understanding, the major known contribution is given by actuation gain noise, contributing on average to about 25% power in the band [60 µHz–1 mHz].

1b. For the analyses at the previous points, we developed tools in the context of multivariate statistical analysis of power spectra, which are completely generalized and not exclusive for LPF analyses. We gave the expression of the generic probability distribution of CPSD matrix, and from that we built a Bayesian analysis for the posterior of derived quantities. Based on that, we developed a decorrelation tool, useful to disentangle coupled sources and analyze their spectra separately.

2. We analyzed LPF's Brownian noise. It is known that the [1 mHz–10 mHz] LPF band was affected by Brownian noise – indeed a proxy for pressure. We showed that the Brownian level in linear acceleration is indeed a good proxy for pressure, and as such can be exploited to analyze pressure evolution and the vacuum environment. This is what we did, finding that the outgassing environment was dominated by a source scaling as $t^{(-0.78\pm0.02)}$, and an activation temperature of (6900 ± 200) K. Moreover, we analyzed LPF's long-term drift from the perspective of vacuum environment, finding that its presence is compatible with a stable pressure unbalance on TM faces.

3. We dedicated plenty of space to the acceleration transients, "glitches" detected on LPF. First, we analyzed the glitch properties and identified two different categories, impacting in different ways LPF's data and potentially impacting in different ways LISA's science. Supported by coincidences with other interferometric channels, we found it highly likely that events in one of the two categories are due to glitches in the interferometric readout.
   We thoroughly analyzed events of the main family, showing up with a rate of 1/day in ordinary runs, increasing during cold runs. We analyzed their amplitude, their impulse, their duration. We finely analyzed the dependence of the glitch rate on temperature, to understand the degree of dependence during the cold runs, and analyzed their correlation to torque time series. In the final chapter, we presented a long discussion about the findings, presenting cross-analyses with other time series, looking for correlations and finding none. As of now, a possibility remains that glitches were caused by gas bursts, though further analyses might be required.

Finally, we presented a long discussion about all the identified low-frequency acceleration features, i.e. the stationary noise, the long-term drift, and the transient glitches, investigating the physical mechanisms ultimately causing them, and highlighting some points to be further analyzed.

Lastly we note that, despite overcoming its requirements, LPF left behind some open questions, and opened some new, unpredicted ones. In particular, questions regarding the physical origin of the measured low-frequency acceleration, which is still largely unmodeled. The results of the presented analyses are not always conclusive, but quite often they tend to exclude possibilities, based on observations.

We are hopeful that some results may be the driver of further studies, to be performed on-ground in view of LISA, to better understand the LPF performance, and, ultimately, to make predictions for LISA.

A List of LPF runs

Table A.1: List of all used LPF runs, label used within this work, legacy LPF index, start and end times, duration, mean electrode housing temperature, actuation authority scheme as in Table A.3. Runs highlighted in blue are cold runs. Runs highlighted in orange refer to long runs used in low-frequency analyses. For those runs, M is the number of periodograms available at 18 µHz.
In the text, we refer to the blue runs in this table as *cold runs*, and to the other ones as *ordinary runs*. The orange ones are referred to as *long-lasting ordinary runs*, or simply *ordinary runs* when unmistakable.
(*) Temperature in cold runs is measured with LTP bay (LCA) sensors, because of the limited working range of EH sensors.

Label	idx	Start time	End time	Days (M)	Temperature [K]	Act.
Mar16A	01	2016-03-01 08:05:00	2016-03-02 23:59:59	1.7	(295.49 ± 0.01)	NOM
Mar16B	02	2016-03-03 15:00:00	2016-03-04 20:59:59	1.2	(295.46 ± 0.01)	RED
Mar16C	03	2016-03-06 18:15:01	2016-03-08 07:59:59	1.6	(295.51 ± 0.03)	RED
Mar16D	04	2016-03-13 16:30:00	2016-03-15 06:59:59	1.6	(295.50 ± 0.03)	RLA
Mar16E	06	2016-03-21 01:59:10	2016-03-26 07:59:44	5.3 (3)	(295.37 ± 0.04)	URLA
Mar16F	08	2016-03-31 08:00:00	2016-04-02 01:59:59	1.7	(295.31 ± 0.01)	RED
Apr16A	09	2016-04-04 00:00:00	2016-04-13 07:59:44	9.3 (6)	(295.30 ± 0.03)	UURLA
Apr16B	12	2016-04-26 08:04:00	2016-04-28 07:59:59	2.0	(295.18 ± 0.01)	UURLA
May16A	13	2016-05-01 08:05:00	2016-05-02 23:54:59	1.7	(295.21 ± 0.04)	UURLA
May16B	14	2016-05-03 08:00:00	2016-05-05 15:29:59	2.3	(295.14 ± 0.02)	UURLA
May16C	17	2016-05-16 00:00:00	2016-05-19 04:59:59	3.2 (1)	(294.97 ± 0.06)	UURLA
May16D	18	2016-05-19 11:30:00	2016-05-21 10:59:59	2.0	(294.92 ± 0.01)	NOM
Jun16A	37	2016-06-15 13:35:00	2016-06-18 07:58:59	2.8	(294.92 ± 0.02)	UURLA
Jun16B	39	2016-06-19 13:00:00	2016-06-24 07:59:59	4.8 (2)	(294.93 ± 0.01)	UURLA
Jul16A	41	2016-07-11 11:40:00	2016-07-12 09:54:59	0.9	(294.00 ± 0.04)	UURLA

Jul16B	42	2016-07-17 12:00:00	2016-07-20 05:59:59	2.7 (1)	(296.62 ± 0.05)	UURLA
Jul16C	43	2016-07-24 17:00:00	2016-07-29 23:59:59	5.3 (3)	(296.50 ± 0.02)	UURLA
Jul16D	44	2016-07-31 11:40:00	2016-08-02 05:59:59	1.8	(296.59 ± 0.03)	UURLA
Aug16A	46	2016-08-12 11:30:00	2016-08-13 07:59:59	0.9	(296.45 ± 0.01)	DURLA
Aug16B	48	2016-08-23 14:00:00	2016-08-27 19:59:59	4.2	(296.45 ± 0.01)	DURLA
Sep16A	49	2016-09-05 11:35:00	2016-09-06 05:04:59	0.7	(296.63 ± 0.03)	DURLA
Sep16B	51	2016-09-11 21:15:00	2016-09-16 05:14:59	4.3	(296.52 ± 0.02)	DURLA
Sep16C	52	2016-09-16 18:36:54	2016-09-17 08:00:39	0.6	(296.56 ± 0.01)	DURLA
Sep16D	53	2016-09-19 02:32:00	2016-09-22 05:55:58	3.1	(296.69 ± 0.05)	UURLA
Sep16E	54	2016-09-28 13:35:00	2016-10-01 07:55:58	2.8 (1)	(296.67 ± 0.01)	UURLA
Oct16A	55	2016-10-04 00:00:00	2016-10-05 07:10:51	1.3	(296.64 ± 0.02)	DURLA
Oct16B	56	2016-10-05 17:30:00	2016-10-08 07:49:59	2.6	(296.3 ± 0.1)	UURLA
Nov16A	58	2016-11-07 21:30:00	2016-11-12 07:59:59	4.4	(296.81 ± 0.03)	UURLA
Nov16B	59	2016-11-16 11:05:00	2016-11-26 07:59:59	9.9 (5)	(296.86 ± 0.04)	UURLA
Dec16A	60	2016-12-02 23:35:00	2016-12-04 16:34:59	1.7	(296.67 ± 0.01)	DURLA
Dec16B	61	2016-12-26 08:01:30	2017-01-13 19:57:59	18.5 (13)	(295.38 ± 0.04)	UURLA
Feb17A	64	2017-02-02 07:00:00	2017-02-02 20:15:58	0.6	(284.66 ± 0.05)	UURLA
Feb17B	66	2017-02-14 01:59:50	2017-02-27 09:53:29	13.3 (9)	(284.72 ± 0.03)	UURLA
Apr17A	69	2017-04-21 14:15:00	2017-04-24 07:44:59	2.7 (1)	(285.88 ± 0.01)	DURLA
May17A	70	2017-05-02 08:00:00	2017-05-03 18:26:59	1.4	$(274.9 \pm 0.6)^*$	UURLA
May17B	71	2017-05-03 23:30:00	2017-05-09 13:59:59	5.6	$(274.05 \pm 0.05)^*$	UURLA
May17C	72	2017-05-10 11:15:17	2017-05-12 12:02:06	2.0	$(273.55 \pm 0.17)^*$	UURLA
May17D	73	2017-05-12 12:02:07	2017-05-15 08:00:58	2.8	$(276.5 \pm 1.4)^*$	UURLA
May17E	74	2017-05-18 18:24:46	2017-05-23 01:59:59	4.3 (2)	(284.19 ± 0.13)	UURLA
Jun17A	75	2017-05-29 13:55:10	2017-06-05 07:59:59	6.8 (4)	(295.78 ± 0.01)	UURLA
Jun17B	76	2017-06-08 12:00:45	2017-06-17 01:42:30	8.6 (5)	(295.91 ± 0.02)	UURLA

Table A.2: Experimental configuration of the 13 noise-only runs used for low-frequency noise estimation. Runs correspond to the ones marked in orange in Table A.1, except for run Apr17A which is a DRS run. The meaning of the numeric labels is explained in the text.

Run	Propellant tank	Thruster branch	Actuation authority	TM alignment	Voltage compensation	Heaters configuration	DRS state
Mar16E	2	A	URLA	1	1	1	OFF
Apr16A	2	A	UURLA	1	1	1	OFF
May16C	2	A	UURLA	1	2	1	OFF
Jun16B	3	A	UURLA	2	3	1	OFF
Jul16B	3	A	UURLA	3	3	2	DIAG
Jul16C	1	A	UURLA	3	3	2	DIAG
Sep16E	3	B	UURLA	3	3	2	DIAG
Nov16B	1	B	UURLA	3	3	3	DIAG
Dec16B	1	B	UURLA	3	4	4	OFF
Feb17B	1	B	UURLA	3	4	5	OFF
May17E	3	A	UURLA	3	4	6	OFF
Jun17A	3	A	UURLA	3	4	1	OFF
Jun17B	3	A	UURLA	3	4	1	OFF

Table A.3: Actuation authority settings

	$F_{x,1}$ pN	$F_{x,2}$ pN	$F_{y,1}$ pN	$F_{y,2}$ pN	$F_{z,1}$ pN	$F_{z,2}$ pN	$N_{\theta,1}$ pNm	$N_{\theta,2}$ pNm	$N_{\eta,1}$ pNm	$N_{\eta,2}$ pNm	$N_{\phi,1}$ pNm	$N_{\phi,2}$ pNm
NOM	0	2200	3670	3670	5820	5820	16.37	16.37	13.32	13.32	10	11
RED	0	600	3670	3670	5820	5820	16.37	16.37	13.32	13.32	3	3
RLA	0	200	3670	3670	5820	5820	16.37	16.37	13.32	13.32	3	3
URLA	0	50	3670	3670	5820	5820	16.37	16.37	13.32	13.32	1.5	1
UURLA	0	50	1000	1000	500	500	4.00	4.00	4.00	4.00	1.5	1
DURLA	0	2200	3670	3670	5820	5820	16.37	16.37	13.32	13.32	10	11

Table A.4: Test mass alignment settings. Setpoint indexes are used in Table A.2.

Angle	Setpoint 1 (Mar16–Jun16)	Setpoint 2 (Jun16)	Setpoint 3 (Jun16–Jun17)
ϕ_1 [µrad]	−59.25	−56.32	−61.2
ϕ_2 [µrad]	−21.35	−33.01	−9.7
η_1 [µrad]	−3.5	−2.14	−4.9
η_2 [µrad]	3.5	10.3	−3.3

Table A.5: DC voltages settings. Configuration indexes are used in Table A.2.

DC Voltage	Config. 1 (Mar16–May16)	Config. 2 (May16)	Config. 3 (May16–Nov16)	Config. 4 (Dec16–Jun17)
$\Delta_{x,1}$ [mV]	0	+48	+24	+24
$\Delta_{x,2}$ [mV]	0	0	0	0
$\Delta_{\phi,1}$ [mV]	0	0	0	+32
$\Delta_{\phi,2}$ [mV]	0	0	0	−116

B | Appendix on spectral estimation

B.1 Generation of random Wishart-distributed samples

Some properties of the complex Wishart distribution [61, 97] can be exploited to define a random sample generator. Let's assume that the matrix $\boldsymbol{W}$ is distributed as a complex Wishart $\mathcal{CW}(\boldsymbol{\Sigma}, K)$. The algorithm we report here allows to re-write the matrix in a form in which all the entries are independently distributed, and then reassemble it to get random samples of the matrix $\boldsymbol{W}$. The key points are the following:

1. The matrix $\boldsymbol{\Sigma}$ is Hermitian and positive-definite. Therefore it can be factorized as $\boldsymbol{\Sigma} = \boldsymbol{D}^H \boldsymbol{D}$, where $\boldsymbol{D}$ is an upper triangular matrix, with the diagonal elements being real and positive. The process is known as Cholesky decomposition.

2. If we Cholesky-decompose $\boldsymbol{\Sigma} = \boldsymbol{D}^H \boldsymbol{D}$, then the matrix $\boldsymbol{V}$ such that $\boldsymbol{W} = \boldsymbol{D}^H \boldsymbol{V} \boldsymbol{D}$ follows a complex Wishart distribution $\mathcal{CW}(\boldsymbol{I}, K)$ [97].

3. Now consider the Cholesky factors $\boldsymbol{T}$ of the matrix $\boldsymbol{V}$, such that $\boldsymbol{V} = \boldsymbol{T}^H \boldsymbol{T}$. It can be demonstrated [55, 97] that the entries of this matrix are independent and follow the following distributions:

   (a) The upper-triangular off-diagonal elements are complex Gaussian variables, i.e., their real and imaginary parts can be independently drawn from zero-mean Gaussian with variance 1/2,

   $$T_{i<j} \sim \mathcal{N}(0, 1/2) + i\mathcal{N}(0, 1/2)$$

   (b) The square of the i-th diagonal element is a sample drawn from a Gamma distribution with shape parameter $K - i + 1$, i.e.

   $$T_{ii}^2 \sim \Gamma(K - i + 1)$$

4. Finally, the complex Wishart distributed matrix $\boldsymbol{W}$ can be recovered, as

   $$\boldsymbol{W} = \boldsymbol{D}^H \boldsymbol{T}^H \boldsymbol{T} \boldsymbol{D} \tag{B.1}$$

Posterior $\mathcal{CW}$ generator, Jeffrey's prior. The same random generator can be employed to generate samples from the posterior distribution of the theoretical

CPSD matrices. As noted in Sec. 3.3.1 on page 49, if a prior can be defined on the entire theoretical CPSD matrix $\boldsymbol{\Sigma}$, the inverse-CPSD $\boldsymbol{Q} = \boldsymbol{\Sigma}^{-1}$ follows in its turn a complex-Wishart distribution. The number of degrees of freedom associated with such distribution depends on the prior choice. Therefore, samples from the posterior distribution can be easily generated using the sampling algorithm to draw samples for $\boldsymbol{Q} \sim \mathcal{CW}(\boldsymbol{W}^{-1}, K)$ (assuming Jeffrey's prior), and then compute $\boldsymbol{\Sigma} = \boldsymbol{Q}^{-1}$ on the samples.

B.2 Calculations for time series decorrelation

B.2.1 Derivation of spectral quantities for time series decorrelation

We modeled noise with the equations in Eq. (3.23). Here we deduce the result in Eq. (3.24) and another useful result, in case only one time series $z(t)$ is contributing:

$$\begin{cases} x(t) &= x_0(t) + \int_{-\infty}^{+\infty} \alpha(t-t')z(t)\,dt' \\ y(t) &= z(t) + n(t) \end{cases} \tag{B.2}$$

We write $x(t)$ instead of Δg, for readability.

- First, let's evaluate the spectral cross-correlation between x and $y(t)$, which would appear as the upper-right element of the CPSD matrix, with the nomenclature of Sec. 3.4. To do it, we first have to evaluate the time cross-correlation, classically defined as $R_{xy}(\tau) = \langle x(t+\tau)\, y(t)\rangle$. It is known that the cross-correlation $S_{xy}(\omega)$ is the Fourier transform of the time cross-correlation function. Developing from this definition,

$$\begin{aligned} R_{xy}(\tau) &= \langle x(t+\tau)\, y(t)\rangle = \\ &= \cancel{R_{x_0 z}(\tau)} + \cancel{R_{x_0 n}(\tau)} + \left\langle \int_{-\infty}^{+\infty} \alpha(t')\, z(t+\tau-t')\, z(t)\, dt' \right\rangle + \\ &\quad + \left\langle \int_{-\infty}^{+\infty} \alpha(t')\, z(t+\tau-t')\, z(t)\, dt' \right\rangle = \\ &= \int_{-\infty}^{+\infty} \alpha(t')\, \langle z(t+\tau-t')\, z(t)\rangle\, dt' + \int_{-\infty}^{+\infty} \alpha(t')\, \langle z(t+\tau-t')\, n(t)\rangle\, dt' = \\ &= \int_{-\infty}^{+\infty} \alpha(t')\, R_{zz}(\tau-t')\, dt' + \cancel{\int_{-\infty}^{+\infty} \alpha(t')\, R_{zn}(\tau-t')\, dt'}, \end{aligned} \tag{B.3}$$

  where some terms cancel out, because of mutual independence between such time series. Calculating the Fourier transform of the result leads to

$$S_{xy}(\omega) = \alpha(\omega) S_{zz}(\omega) \tag{B.4}$$

- Analogous reasoning can be applied to the auto-correlation $R_{xx}(\tau)$, leading to

$$S_{xx}(\omega) = S_{x_0 x_0}(\omega) + |\alpha|^2(\omega) S_{zz}(\omega) \tag{B.5}$$

B.2.2 Derivation of spectral quantities for time series decorrelation

In Sec. 3.4.2, we state that spectral quantities (specifically, the susceptibilities and the residual PSD) can be more easily expressed as elements of the inverse-CPSD matrix $\boldsymbol{Q}$. In particular, we assume that the theoretical PSD matrix $\boldsymbol{\Sigma}$ can be written as

$$\boldsymbol{\Sigma} = \left(\begin{array}{c|c} S_{\Delta g_0} + \boldsymbol{\alpha} \cdot \boldsymbol{S}_{yy} \cdot \boldsymbol{\alpha}^H & \boldsymbol{\alpha} \cdot \boldsymbol{S}_{yy} \\ \hline \boldsymbol{S}_{yy}^H \cdot \boldsymbol{\alpha}^H & \boldsymbol{S}_{yy} \end{array}\right) \tag{B.6}$$

where $\boldsymbol{\alpha}$ are the complex susceptibilities, $\boldsymbol{S}_{yy}$ is the CPSD matrix of the signals y_i, and $S_{\Delta g_0}$ is the PSD of the so-called residual, i.e. the component of $S_{\Delta g}$ which does not correlate with the signals y_i.
The following equation is a well-known result of block matrix inversion. In particular, if the matrix $\boldsymbol{\Sigma}$ can be partitioned into four blocks as

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix}, \tag{B.7}$$

where $\boldsymbol{A}$ and $\boldsymbol{D}$ are square blocks, it follows that the inverse $\boldsymbol{Q} = \boldsymbol{\Sigma}^{-1}$ reads

$$\boldsymbol{Q} = \begin{pmatrix} (\boldsymbol{A} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C})^{-1} & -(\boldsymbol{A} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C})^{-1}\boldsymbol{B}\boldsymbol{D}^{-1} \\ -\boldsymbol{D}^{-1}\boldsymbol{C}(\boldsymbol{A} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C})^{-1} & \boldsymbol{D}^{-1} + \boldsymbol{D}^{-1}\boldsymbol{C}(\boldsymbol{A} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C})^{-1}\boldsymbol{B}\boldsymbol{D}^{-1} \end{pmatrix} \tag{B.8}$$

From this equation, and from the previous Eq. (B.6), some results follow directly:

1. The complex susceptibility vector, $\boldsymbol{\alpha}$, can be written explicitly from terms of the inverse CPSD $\boldsymbol{Q}$, as

$$\frac{\boldsymbol{Q}_{1,j+1}}{Q_{11}} = \frac{-(\boldsymbol{A} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C})^{-1}\boldsymbol{B}\boldsymbol{D}^{-1}}{(\boldsymbol{A} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C})^{-1}} = -\boldsymbol{B}\boldsymbol{D}^{-1} = \boldsymbol{\alpha}\,\boldsymbol{S}_{yy}\boldsymbol{S}_{yy}^{-1} = \boldsymbol{\alpha} \tag{B.9}$$

2. The residual noise can be written as

$$\begin{aligned} S_{\Delta g_0} &= S_{\Delta g} - \boldsymbol{\alpha} \cdot \boldsymbol{S}_{yy} \cdot \boldsymbol{\alpha}^H = S_{\Delta g} - \boldsymbol{S}_{\Delta gy} \cdot \boldsymbol{S}_{yy}^{-1} \cdot \boldsymbol{S}_{\Delta gy}^H = \\ &= (\boldsymbol{A} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C}) = 1/Q_{11} \end{aligned} \tag{B.10}$$

3. The residual noise coincides with the Schur complement of the matrix $\boldsymbol{S}_{yy}$ in $\boldsymbol{\Sigma}$.

B.2.3 Marginalization of residual PSD posterior

We prove the result in Eq. (3.53), where we state that, under the assumption that susceptibilities are real-valued, and Jeffrey's prior is assumed on the residual PSD $S_{\Delta g_0}$, the single-frequency residual follows an inverse-gamma distribution

$$S_{\Delta g_0} \sim \mathrm{invGamma}\left(M - r/2, M\,\mathrm{Re}\boldsymbol{\Pi}/\mathrm{Re}\boldsymbol{P}_{yy}\right). \tag{B.11}$$

First of all, let's integrate the likelihood function Eq. (3.49) with respect to $\boldsymbol{\alpha}$:

$$p(\{S_{\Delta g_0}\}|\{\boldsymbol{W}\}) \propto \frac{1}{S_{\Delta g_0}^M} \int_{\boldsymbol{\alpha}} \exp\left[-\frac{\boldsymbol{\eta}\cdot\boldsymbol{W}\cdot\boldsymbol{\eta}^H}{S_{\Delta g_0}}\right] d\boldsymbol{\alpha} \times \text{prior}\left(S_{\Delta g_0}\right) \tag{B.12}$$

Let's elaborate on the exponential term. First, as $\boldsymbol{\eta}$ in real, $\boldsymbol{\eta}\cdot\boldsymbol{W}\cdot\boldsymbol{\eta}^H = \boldsymbol{\eta}\cdot\boldsymbol{V}\cdot\boldsymbol{\eta}^T$, where $\boldsymbol{V}$ is the real part $\boldsymbol{V} = \text{Re}\boldsymbol{W}$.
The exponent can be rewritten:

$$\begin{aligned}
\boldsymbol{\eta}\cdot\boldsymbol{V}\cdot\boldsymbol{\eta}^T &= (-1,\boldsymbol{\alpha})\cdot\boldsymbol{V}\cdot(-1,\boldsymbol{\alpha})^T = V_{11} - \boldsymbol{V}_{1y}\cdot\boldsymbol{\alpha}^T - \boldsymbol{\alpha}\cdot\boldsymbol{V}_{1y}^T + \boldsymbol{\alpha}\cdot\boldsymbol{V}_{yy}\cdot\boldsymbol{\alpha}^T = \\
&= V_{11} - \boldsymbol{V}_{1y}\boldsymbol{V}_{yy}^{-1}\boldsymbol{V}_{1y}^T + \boldsymbol{V}_{1y}\boldsymbol{V}_{yy}^{-1}\boldsymbol{V}_{1y}^T - \boldsymbol{V}_{1y}\cdot\boldsymbol{\alpha}^T - \boldsymbol{\alpha}\cdot\boldsymbol{V}_{1y}^T + \boldsymbol{\alpha}\cdot\boldsymbol{V}_{yy}\cdot\boldsymbol{\alpha}^T = \\
&= \left(V_{11} - \boldsymbol{V}_{1y}\boldsymbol{V}_{yy}^{-1}\boldsymbol{V}_{1y}^T\right) + \left(\boldsymbol{V}_{1y}\boldsymbol{V}_{yy}^{-1} - \boldsymbol{\alpha}\right)\cdot\boldsymbol{V}_{yy}\cdot\left(\boldsymbol{V}_{1y}\boldsymbol{V}_{yy}^{-1} - \boldsymbol{\alpha}\right)^T = \\
&= (\boldsymbol{V}/\boldsymbol{V}_{yy}) + \left(\boldsymbol{V}_{1y}\boldsymbol{V}_{yy}^{-1} - \boldsymbol{\alpha}\right)\cdot\boldsymbol{V}_{yy}\cdot\left(\boldsymbol{V}_{1y}\boldsymbol{V}_{yy}^{-1} - \boldsymbol{\alpha}\right)^T,
\end{aligned} \tag{B.13}$$

where the first term is the Schur complement. Hence we can rewrite the integral as

$$\begin{aligned}
&\int_{\boldsymbol{\alpha}} \exp\left[-\frac{\boldsymbol{\eta}\cdot\boldsymbol{W}\cdot\boldsymbol{\eta}^H}{S_{\Delta g_0}}\right] d\boldsymbol{\alpha} = \\
&\exp\left[-\frac{\boldsymbol{V}/\boldsymbol{V}_{yy}}{S_{\Delta g_0}}\right] \int_{\boldsymbol{\alpha}} \exp\left[-\frac{\left(\boldsymbol{V}_{1y}\boldsymbol{V}_{yy}^{-1} - \boldsymbol{\alpha}\right)\cdot\boldsymbol{V}_{yy}\cdot\left(\boldsymbol{V}_{1y}\boldsymbol{V}_{yy}^{-1} - \boldsymbol{\alpha}\right)^T}{S_{\Delta g_0}}\right] d\boldsymbol{\alpha} \propto \\
&\propto S_{\Delta g_0}^{r/2} \exp\left[-\frac{\boldsymbol{V}/\boldsymbol{V}_{yy}}{S_{\Delta g_0}}\right]
\end{aligned} \tag{B.14}$$

where we have used the Gaussian integral. Hence the posterior likelihood is

$$p(\{S_{\Delta g_0}\}|\{\boldsymbol{W}\}) \propto \frac{1}{S_{\Delta g_0}^{M-r/2}} \exp\left[-\frac{\boldsymbol{V}/\boldsymbol{V}_{yy}}{S_{\Delta g_0}}\right] \times \text{prior}\left(S_{\Delta g_0}\right), \tag{B.15}$$

which, with Jeffrey's prior, is the result we wanted to prove.

B.3 Generation of time series with given noise

For algorithm testing purposes (Sec. 3.6.3), we have been in need of creating synthetic time series, lasting as long as the Feb17B run, to run the tests in a controlled environment. These are the requirements:

1. The main measurement $\Delta g(t)$ should be corrupted with two time series $A(t)$ and $B(t)$, via two coupling susceptibilities α_a and α_b.
2. The two signals A and B should have a certain degree of correlation with one another.
3. The two signals A and B should be affected by additional independent readout noises, $n_A(t)$ and $n_B(t)$.
4. For better resembling the LPF case, Δg_0 should have a $\sim 1/f^2 + \text{const.}$ power spectrum; $a(t)$ and $b(t)$ should have a $\sim 1/f$ and $\sim 1/f^2$ power spectrum. We simulate $n_A(t)$ and $n_B(t)$ with the same spectra as $a(t)$ and $b(t)$.

To satisfy these requirements, we independently generate $\{\Delta g_0, a, b, n_A, n_B\}$ with the given power spectra, and mix them as follows:

$$\begin{pmatrix} \Delta g \\ A \\ B \end{pmatrix} = \begin{pmatrix} 1 & \alpha_a & \alpha_b & 0 & 0 \\ 0 & 1 & \mu & \eta & 0 \\ 0 & \nu & 1 & 0 & \eta \end{pmatrix} \begin{pmatrix} \Delta g_0 \\ a \\ b \\ n_A \\ n_B \end{pmatrix} \tag{B.16}$$

Here, we arbitrarily choose $\mu = 0.4$ and $\nu = 0.3$. We choose $\{\alpha_a, \alpha_b\} = 2.5, 4.0$ in test #2.1 and $\{\alpha_a, \alpha_b\} = 0.2, 0.5$ in test #2.2. The spectra $\alpha_a^2 S_a$ and $\alpha_b S_b$ are depicted in text in Fig. 3.8 on page 70. η varies from 0 upwards, to change the impact of the readout noise on A and B.

We note that, from this definition, the expected susceptibilities of A and B differ from α_a and α_b:

$$\begin{pmatrix} \alpha_A & \alpha_B \end{pmatrix} = \begin{pmatrix} \dfrac{\alpha_a - \nu\alpha_b}{1-\mu\nu} & \dfrac{\alpha_b - \mu\alpha_a}{1-\mu\nu} \end{pmatrix} \tag{B.17}$$

C | Appendix on LPF glitches

C.1 Estimation of impulse-carrying glitch parameter errors

We have already noted, in Sec. 7.2.3 on page 143, that the fitting procedure of ICG glitches to the templates in Eqs. (7.1, 7.2, 7.5) is non-optimal and non-linear, because of the presence of red noise at low frequencies. In particular, the red noise is present at the characteristic frequencies of the events themselves, $\sim 1/\tau$. It can be proved that the least squares fit coincides with the Wiener optimal filter only in the case of frequency-independent white noise. Hence, we estimated the fitting parameter covariance from the Cramér-Rao bound [98], assuming the noise is Gaussian, and using the measured PSD of the residuals. For glitches in ordinary runs, we have checked that parameters returned by an optimal filter procedure agree, within the estimated uncertainty, with those found with our non-optimal least squares method.

To this aim, once a glitch had been identified and fitted, we have: expanded the fitting function around the best-fit parameter values, up to linear terms in the fitting parameters; applied the optimal linear filter method for multi-component signals [99]; calculated the fitting amplitudes; propagated these results back to that of the original fitting parameters. Both procedures' results agree with the uncertainty estimated from the Cramér-Rao bound, except for a few outliers.

The optimal linear filter procedure works as follows (regarding single-τ glitches, but it can be extended to multi-τ):

1. We focus on the glitch Fourier template, $h_1(\omega)$ in Eq. (7.3), which depends on Δv, t_0, and τ. We re-define the definition with respect to the variables $q = \Delta v\, t_0$, $r = \Delta v\, \tau$.

2. We linearly expand the Fourier template about the fit parameters, defining the gradient

$$\boldsymbol{\xi}(\omega) = \boldsymbol{\nabla}_{(\Delta v, \Delta v t_0, \Delta v \tau)} h_1(\omega) = \left(\frac{e^{-i\omega t_0}}{(1+i\omega\tau)^2}, -\frac{i\omega\, e^{-i\omega t_0}}{(1+i\omega\tau)^2}, -\frac{2i\omega\, e^{-i\omega t_0}}{(1+i\omega\tau)^3} \right)$$

3. The gradient defines the signal matrix $\boldsymbol{\Xi}(\omega)$, whose elements are defined by $\Xi_{ij}(\omega) = \xi_i(\omega)\xi_j^*(\omega)$. Similarly to the standard 1-dimensional Wiener theory,

the inverse optimal variance matrix is given by

$$\boldsymbol{V}^{-1} = \int_{-\infty}^{+\infty} \frac{\boldsymbol{\Xi}(\omega)}{S_{\text{one-sided}}(\omega)}\, d\omega = 4 \int_{0}^{+\infty} \frac{\mathrm{Re}\,[\boldsymbol{\Xi}(f)]}{S(f)}\, df$$

4. If $\boldsymbol{J}$ is the Jacobian of the transformation $(\Delta v, t_0, \tau) \to (\Delta v, q, r)$, then the matrix $\boldsymbol{C} = \boldsymbol{J}^{-1}\boldsymbol{V}\boldsymbol{J}^{-T}$ is the optimal covariance matrix, with respect to the parameters $(\Delta v, t_0, \tau)$, which we use for error estimation.

5. From $\boldsymbol{C}$, it is straightforward to see that its first and last diagonal elements are the estimated variances of Δv and τ.
   The variance of the duration Δ can be found easily in case the glitch follows the template $h_1(t)$. In the $h_2(t)$ case, the variance is found by standard error propagation, following the definition $\Delta = \Delta(\tau_1, \tau_2)$.
   The variance $\sigma^2_{\Delta v}$ of the amplitude Δv also defines the signal-to-noise ratio, SNR, of the glitch. If an event's amplitude is Δv, then its SNR is defined as $\mathrm{SNR} = |\Delta v|/\sigma_{\Delta v}$.

6. Within the same theory and the linear approximation, also the parameter values can be recovered, evaluating their deviation from the fit values

$$\delta(\Delta v, \Delta v\, t_0, \Delta v\, \tau) = 4 \int_{0}^{+\infty} \frac{\boldsymbol{V} \cdot \boldsymbol{\nabla} h_1(f)\, x(f)}{S(f)}\, df,$$

   where $x(f)$ is the Fourier transform of the data stretch. Nevertheless, we were not able to apply the optimal procedure to the cold runs data because of their above-mentioned complexity, due to overlap of events. Thus, for consistency, we used for both ordinary and cold runs the parameter values resulting from the non-optimal least squares procedure.

Our error estimates represent a lower bound. In particular, given the length of the low pass filter we use for glitch identification, the uncertainty on Δ for the shortest glitches, $\Delta \lesssim 30\,\mathrm{s}$, may be significantly underestimated. However, we stress that none of the results depends critically on the accuracy of such an uncertainty estimate, as parameter fluctuations within the glitch population are significantly larger than their uncertainties.

Standard, one-dimensional case. We note some differences with the standard one-dimensional case of Wiener filters, which is usually used in template searches. The difference is that, in our case, the shape parameters τ_1 and τ_2 are not known. If they were, the vector $\boldsymbol{\xi}(\omega)$ would just reduce to its first component $\xi_1(\omega)$, and the optimal standard deviation would converge to the classically known

$$\sigma_{\Delta v} = \left[4 \int_{0}^{+\infty} \frac{|\xi_1(f)|^2}{S(f)} df \right]^{-1/2},$$

which is the commonly used form. The SNR would thus be defined slightly differently, resulting generally in a higher value. This is understandable, since an initial better

knowledge of the template one is looking for should lower the detection threshold.

C.2 Multi-dimensional matched-filtering for noise-correlated detectors

We consider a generic array of detectors, indexed by i. Each of these detectors is characterized by a known background noise PSD, due both to the common noise and to the different readout noises. The presence of a common noise induces a non-zero cross-spectral density between the noise-only time series, so it is convenient to define a cross-spectral matrix $\mathbf{\Sigma}(\omega)$ (see Chapter 3).

Let's assume that the i-th detector measures the $x_i(t)$ time series. Let's also assume that the transient signal occurs at time t_0, matching a known template $h_i(t)$ with an amplitude A_i. This signal adds to the detectors' noises $n_i(t)$, so that

$$x_i(t) = A_i h_i(t) + n_i(t) \tag{C.1}$$

In our case, the signals h_i have the same shape h and amplitude A in both detectors. As said, the detector noises $n_i(t)$ might cross-correlate.

Extending the usual Wiener filter calculations, we prove that the optimal multi-detector filter, in the frequency domain, reads:

$$\boldsymbol{H}(f) = \frac{\boldsymbol{h}(f) \cdot \mathbf{\Sigma}^{-1}(f)}{\sigma_A^2}, \qquad \sigma_{\hat{A}}^2 = \left(4 \int_0^{+\infty} \boldsymbol{h}(f) \cdot \mathbf{\Sigma}^{-1}(f) \cdot \boldsymbol{h}^H(f) df \right)^{-1} \tag{C.2}$$

here, $\sigma_{\hat{A}}^2$ is the variance associated with the filter, corresponding to the optimal variance of the Cramér-Rao bound. When applied to data $\boldsymbol{x}(f)$, the filter yields the estimate of the amplitude $\hat{A}$. The signal-to-noise ratio follows then the usual definition, $\mathrm{SNR} = \hat{A}/\sigma_{\hat{A}}^2$.

The filter defined in Eq. (C.2) is greater as the intrinsic noise of a detector is lower, hence privileging those detectors with low noise. We note that, in case one of the detector is much noisier than the others, it has virtually no influence in Eq. (C.2), practically resulting excluded from the analyses.

D | Molecular simulations of LPF vacuum environment

To perform simulations of LPF's outgassing environment and pressure properties, we employed `Molflow+`, a stand-alone software developed at CERN [87, 88]. On Molflow's website, it is possible to find extensive documentation about the algorithm behind it; however, we summarize the main details, important to our purposes.

Molflow is a Monte Carlo simulator, intended to calculate pressure profiles and conductances in ultra-high vacuum environments, given arbitrary custom geometries. The assumption of ultra-high vacuum is key, because interactions need to be modeled with the molecular free flow regime. If the pressure is relatively high (usually, a reference threshold of 10^{-3} mbar is set), interactions between molecules are relevant. So relevant, that inter-molecular forces might become dominant; in that case, the problem of pressure evolution would enter the field of fluid dynamics and the viscous regime, with all its known difficulties, and non-linear differential equations. In our case, pressure is low, of the order of 10^{-7}–10^{-8} mbar, and molecules move freely without interacting with each other. The molecular free path, indeed, is $>$ km. This is an enormous advantage from the computational point of view, as it allows to simulate and track molecules one by one in parallel, without caring about inter-molecular forces and interactions, but rather caring only about the interactions with the surfaces.

In the past, simulations of LPF's vacuum environment were performed [3, 22], often with simplified geometries with a few surfaces representing just the inner EH walls, electrodes, laser holes. Molflow offers a more advanced setup, allowing to import custom CAD geometries. The software is well optimized, so that it can simulate systems of a few $\sim 10^5$ faces, within reasonable computational time.
In Fig. D.1, we show a snapshot of a simulation.

D.1 Algorithm layout

The idea behind Molflow is the following: a single core works on a single particle, from its emission, until it is pumped out of the system. The aforementioned particle, in its journey toward the pumping ducts, or toward any sticking element, hits the surrounding surfaces, transferring momentum and, hence, giving pressure. Many particles can be simulated, even parallelizing more processes in a multi-core system. The reason for this is always the molecular regime: pressure (simulated), generated by a number of simulated (virtual) particles is proportional to real pressure, generated by a flow of real particles. Simulating a higher number of particles, of course, allows

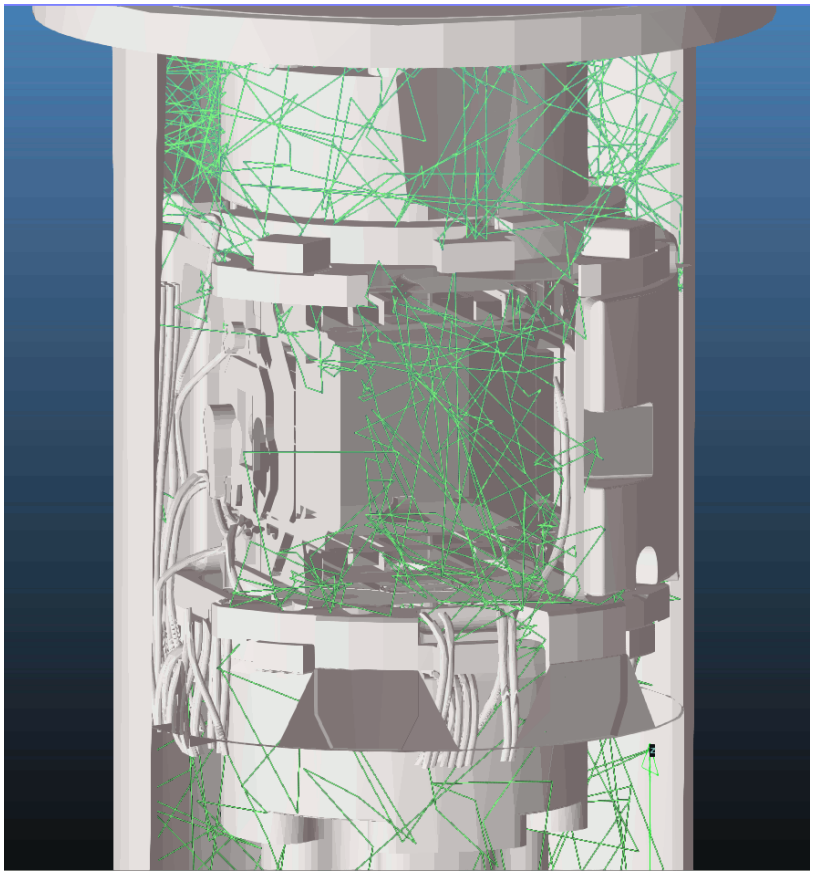


Figure D.1: Snapshot of a simulation of a particle emitted in the GRS, by the IBM, with our simplified version of the IS geometry. The IBM is visible on the right. In the center, the TM and its EH, with many cables and harnesses, and their connectors. Only a few hundred hits are shown. The particle is finally absorbed by the venting duct, not visible. [Geometry courtesy, OHB Italia]

to improve the precision of the estimate.
The simulation proceeds with the following outline:

- **Emission.** Given the settings chosen by the user, a starting location is chosen, according to the local outgassing flux. The particle direction is randomly chosen, accordingly to the standardly implemented Knudsen's rule, or cosine-law rule, which states that the probability of being emitted in any direction is independent of the azimuth angle ϕ, but it only depends on the polar angle θ. The maximum probability, hence, is to be emitted perpendicularly to the surface. Also, the particle speed is chosen accordingly to the proper Maxwell distribution[1], depending on the molecule mass and the surface temperature. For example, the typical velocity of a water molecule emitted by a surface at 293 K, is 520 m/s.

- **Traveling.** Molflow does not use a time-evolution algorithm. Rather, it implements a ray-tracing method to find the next collision with another facet, exploiting the fact that molecules do not interact with each other. The time the ray-tracing method takes, does not depend on the time it takes the particle in its journey. In our case, this saves computational time. Once a surface is reached, the particle could be pumped away, or rebound.

  - Each facet is provided with a sticking factor, which is the probability that the surface has to remove the particle from the system. Obviously, pumps are treated as 100% sticking surfaces. When a surface is hit, a random number is extracted to choose if it is absorbing the particles or not. If the particle is absorbed, then it is simply removed from the system, and the core proceeds with another particle.
  - If the particle is rebound, a new speed and direction are chosen. The key assumption is that particles thermalize to surfaces during collisions. Collisions last usually a negligible time with respect to the molecule travel times (see later, sojourn time), but those times are enough for the particle to be re-emitted according to the cosine law and Maxwell's law, with the temperature of the hit surface.

- **Absorption.** The particle journey ends with an absorption, by a sticking surface or pump.

- **Physical quantities.** Physical quantities are calculated and recorded every time a particle hits a surface. In the standard case, Molflow is intended to be used to estimate pressure in a particular location, which is evaluated starting by the momentum transfer in the direction perpendicular to the surface. In LPF's case, we use a slightly upgraded version of the Molflow software. The relevant quantity is indeed the force on the TM along the x axis, hence we need to account for all momentum transfers along the x direction, careless of the orientation of the surface. Therefore, for each virtual particle, we record the x component of

[1] To be precise, the software also applies a correction accounting for the fact that faster particles get to distant surfaces in shorter times than the others.

the molecule impulse variation during the collision, $\Delta I_x = I_{x_{\text{out}}} - I_{x_{\text{in}}}$. From this, force reads:

$$F_x = \frac{\sum dI_{x,\text{real}}}{dt} = \sum \Delta I_{x,\text{virtual}} \frac{dN_{\text{real}}/dt}{N_{\text{virtual}}} \tag{D.1}$$

Another similar-fashioned quantity, available from the latest software versions and unimplemented in the presented analyses, is the torque on the TM.

- **Time-dependent simulations.** Simulations might either be steady-state, or time-dependent. In a steady-state simulation, a constant molecular flow is simulated, representing a constant outgassing rate $\dot{Q}$. A constant flow of molecules dN/dt yields a constant pressure, or force, on the region of interest. Otherwise, it is possible to simulate a time-dependent molecular flow, yielding time-dependent physical quantities. In this case, we argue that the best time-saving option is to simulate the response to a delta-like outgassing impulse $\delta(t)$, which from the numerical point of view can just be approximated as a simple step. Afterward, once having the impulse response $F_\delta(t)$, representing the time evolution of TM force due to the outgassing pulse, it is straightforward to convolve it with the desired outgassing profile $\dot{Q}(t)$, to get the time profile of the force $F(t)$.

  $$F(t) = (\dot{Q} * F_\delta)(t) \tag{D.2}$$

  Molflow also offers the possibility to look at the evolution of the output quantities during simulations, to be used as convergence checks. Usually, a few 100×10^3 simulated virtual particles (corresponding to some 100×10^6 total hits) are enough to draw conclusions, but it really depends on the needed precision and how weak is the investigated phenomenon.

- **Sojourn time.** In describing the molecule traveling, we made a strong assumption, saying that the molecule residence time on the surface is negligible with respect to the travel time. This, in general, is not true. As documented in [100], molecules interact with surface walls, and get adsorbed with binding energies which depend on both the traveling molecule, the absorbing surface elements, and the bonding type. The higher the binding energy, the lower the molecule escape probability will be, hence the higher the residence time on the wall will be. The residence time is usually known as the *sojourn time*. The mean sojourn time can be as low as 10^{-12} s, which holds for instance for helium on metals, to 10^2 s, for chemisorbed molecular hydrogen. It is hard to say what is the expected sojourn time in the LPF case, but it can hardly be greater than a few seconds, given that the most probable outgassing gas is water, from metals (Mo, Ti, W). Once absorbed with mean sojourn time τ, a molecule will escape at a time t, governed by an exponential probability $\propto \exp(-t/\tau)$.
  Molflow allows for simple simulations of the sojourn time: once absorbed, the time of re-emission is randomly chosen from the proper exponential distribution.

D.2 Usage in the LPF case

In the LPF case, we use Molflow in both constant-flow and time-dependent modes. A snapshot of a simulation is shown in Fig. D.1. The geometry is representative of the real flight-model IS (courtesy OHB Italia S.p.A), with "radially increasing" simplification: the TM, as well as the inner EH, is not simplified. The outer EH faces are slightly simplified: screws, bolts, nuts are removed, fibers are removed, and their holes are filled. The structure is more simplified, chamfers are removed and rectified. Outer components, such as mechanical parts, motors, etc. are removed. The vent valve/duct is simply represented as a sticking surface. The model is meant to be representative of the inner TM region, not of the venting parts. Hence, surfaces are re-meshed, so that the number of facets is reduced to a few hundred thousand.
In Fig. D.2, we show a histogram of the permanence time within the IS of a water molecule which has been emitted from the inner EH, assuming no sojourn time on the walls, and assuming that only the vent duct is absorbing.

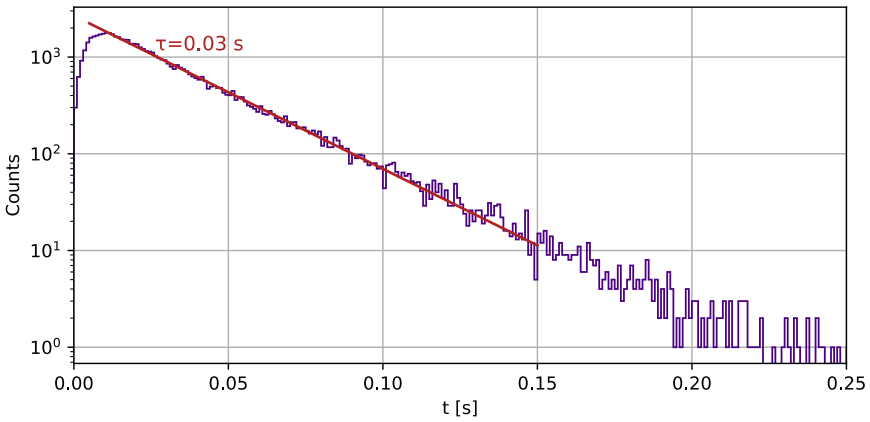


Figure D.2: Histogram of the permanence time of water molecules emitted by the inner EH, before absorption by the venting duct, assuming no sojourn time on the walls. The decay time of the exponential pattern, represented as a red solid line, is $\tau = 0.03\,\mathrm{s}$.

Acknowledgments

This thesis started with a quote, on which I would like to spend a few words. In trying to figure out which sources contributed to observations, we went over a large number of physical mechanisms, possibly and presumably leading to the observations we wanted to explain. We dug into the finest details of LPF measurements, looking for correlations, and any possible clues pointing to yet unmodeled mechanisms. Some of these mechanisms, at first sounded plausible. Sometimes though, it was frustrating to understand that the effect we were considering was of negligible impact; even more so, if we had been analyzing it for a long time.
I believe this is just an example of science itself, showing the importance of going deeper and deeper into concepts. Even the slightest inconsistency should not be left unnoticed, and actively searching for it is a task of research, even though it might sound self-defeating.

Many people contributed to the development of this work, either directly or indirectly, aware or unaware of it, during these 3+ years of PhD.

First of all, I'd like to thank my supervisor Stefano Vitale, for his boundless passion for science, for his constant dedication to the LPF and LISA missions, and for teaching me how to persevere in research. I would also like to thank Bill Weber and Rita Dolesi, for their genuine curiosity about nature, and for their determination in making things work. I thank them, and Stefano, for being able to create a group that is based on cooperation, instead of competition. Working in a friendly scientific environment that is driven by curiosity is something rare, it should not be taken for granted.

Immense thanks go to Eleonora Castelli, awesome officemate and friend. She introduced me to the magical world of the LPF mission, to the unwritten knowledge of the LPF operations, and shared with me the tools and tricks of the LPF data analysis and LTPDA.

I'd also like to thank Carlo Zanoni: I had no idea how to make the beautiful renderings he made for the first chapters of this thesis. Apart from this, I'd really like to thank him for his insights and discussions on a scientist's career, and for the valuable opportunities. I also thank Valerio Ferroni and Daniele Vetrugno for their insights into the unwritten/unspoken details of LPF, and all the LISA PhDs (mostly, former PhDs) Martina Muratore, Davide Dal Bosco,

and Vittorio Chiavegato, and the rest of the UTN LISA group, for being so united in achieving a common objective.
I am also grateful to the members of the LPF Collaboration, the LISA Consortium, and the LISA Artifacts group, who contributed to improving the analyses on LPF glitches. Furthermore, the members from the Vacuum-surfaces-coatings group at CERN, and the AEI Institute Hannover, for the enlightening discussions on, respectively, vacuum science and LPF's interferometer.

A PhD is not just research. It is also people, and I learned that by doing it. Thanks to Sergio. I really owe him a lot, for his resoluteness, and his willingness to give advice and support whenever needed. And for being my first ski instructor, which is not a thing to be underestimated in Trento.
I'd also like to thank Daniele, Piero, and Alberto. Being all day in Povo would be much more boring without them. And thanks to Philipp for the squash matches, though I had to let him down at one point.

This PhD time wouldn't have been the same without ADI, the association of PhD candidates and graduates in Italy. Motivated people, determined in struggling for researchers' rights, and willing to devote their time to help others; it has been an honor to take part in the regrowth of the Trento branch, I really hope that the atmosphere it has right now can last for a long time.
ADI also made me meet my dearest friends in Trento. These are the people who really made a difference in this PhD life, it definitely wouldn't have been the same without them. Thanks to Samuele, Alice, Charlotte, Evelyn, Nishal, Marija, Davide, Stefan, Adelaide, and many many others.

As the last lines of these acknowledgments, and of this thesis, I'd spend some words for the people who have always been there, my friends from Milan, all my family, and my parents. Their unconditional support is a solid certainty, a fixed point I can always rely on.

Bibliography

[1] M. Armano et al. "Sub-Femto-g Free Fall for Space-Based Gravitational Wave Observatories: LISA Pathfinder Results". In: *Phys. Rev. Lett.* 116 (23 2016), p. 231101. DOI: 10.1103/PhysRevLett.116.231101.

[2] M. Armano et al. "Beyond the Required LISA Free-Fall Performance: New LISA Pathfinder Results down to 20 μHz". In: *Phys. Rev. Lett.* 120 (6 Feb. 2018), p. 061101. DOI: 10.1103/PhysRevLett.120.061101.

[3] Eleonora Castelli. "LISA Pathfinder noise performance results: disturbances in the sub-mHz frequency band and projection to LISA". PhD thesis. University of Trento, 2020. DOI: 10.15168/11572_254388.

[4] M. Armano et al. "Transient acceleration events in LISA Pathfinder data: Properties and possible physical origin". In: *Phys. Rev. D* 106 (6 Sept. 2022), p. 062001. DOI: 10.1103/PhysRevD.106.062001.

[5] LISA Pathfinder collaboration. "In-depth analysis of LISA Pathfinder performance results: time evolution, noise projection, physical models and implications for LISA". In preparation.

[6] Amaro-Seoane et al. *Laser Interferometer Space Antenna*. 2017. DOI: 10.48550/arXiv.1702.00786.

[7] LISA Collaboration. *The Gravitational Universe*. White Paper. European Space Agency, 2014.

[8] LISA Instrument Group. *LISA Payload Definition Document*. Technical note ESA-L3-EST-INST-DD-001. LISA Consortium, 2018.

[9] Charles W. Misner, Kip S. Thorne, John Archibald Wheeler, and David Kaiser. *Gravitation*. OCLC: on1006427790. Princeton, N.J: Princeton University Press, 2017. ISBN: 9780691177793.

[10] The LIGO Scientific Collaboration. "Advanced LIGO". In: *Classical and Quantum Gravity* 32.7 (Mar. 2015), p. 074001. DOI: 10.1088/0264-9381/32/7/074001.

[11] *LISA Assessment Study Report*. Tech. rep. ESA/SRE(2001)3. ESA, 2011.

[12] LISA Instrument Group. *LISA Performance Model and Error Budget*. Technical note LISA-LCST-INST-TN-003. LISA Consortium, 2020.

[13] M. Armano et al. “Sensor Noise in LISA Pathfinder: In-Flight Performance of the Optical Test Mass Readout”. In: *Phys. Rev. Lett.* 126 (13 Apr. 2021), p. 131103.

[14] F. Antonucci et al. “From laboratory experiments to LISA Pathfinder: achieving LISA geodesic motion”. In: *Classical and Quantum Gravity* 28.9 (2011), p. 094002.

[15] P. Sarra. *TM FM3 Magnetic Moment Measurement Report*. Tech. rep. S2-CGS-TR-3030. OHB-CGS, 2012.

[16] LISA Pathfinder Collaboration, M. Armano, et al. “Capacitive sensing of test mass motion with nanometer precision over millimeter-wide sensing gaps for space-borne gravitational reference sensors”. In: *Physical Review D* 96.6 (Sept. 2017), p. 062004. DOI: 10.1103/PhysRevD.96.062004.

[17] Pierre Touboul and Manuel Rodrigues. “The MICROSCOPE space mission”. In: *Classical and Quantum Gravity* 18.13 (July 2001), p. 2487. DOI: 10.1088/0264-9381/18/13/311.

[18] Pierre Touboul et al. “MICROSCOPE Mission: Final Results of the Test of the Equivalence Principle”. In: *Phys. Rev. Lett.* 129 (12 Sept. 2022), p. 121102. DOI: 10.1103/PhysRevLett.129.121102.

[19] B. D. Tapley, S. Bettadpur, M. Watkins, and C. Reigber. “The gravity recovery and climate experiment: Mission overview and early results”. In: *Geophysical Research Letters* 31.9 (2004). DOI: https://doi.org/10.1029/2004GL019920.

[20] William J. Weber et al. “Position sensors for flight testing of LISA drag-free control”. In: ed. by Mike Cruise and Peter Saulson. Waikoloa, Hawai’i, United States, Feb. 2003, p. 31. DOI: 10.1117/12.458564.

[21] R Dolesi et al. “Gravitational sensor for LISA and its technology demonstration mission”. In: *Classical and Quantum Gravity* 20.10 (May 2003), S99–S108. ISSN: 0264-9381, 1361-6382. DOI: 10.1088/0264-9381/20/10/312.

[22] A. Cavalleri et al. “Increased Brownian Force Noise from Molecular Impacts in a Constrained Volume”. In: *Phys. Rev. Lett.* 103 (14 Sept. 2009), p. 140601.

[23] LISA Pathfinder collaboration. “Nanonewton electrostatic actuators for femtoNewton small-force measurements: performance of the LISA Pathfinder electrostatic force actuation system”. In preparation.

[24] M. Armano et al. “Calibrating the system dynamics of LISA Pathfinder”. In: *Phys. Rev. D* 97 (12 June 2018), p. 122002.

[25] W.J. Weber et al. “Possibilities for measurement and compensation of stray DC electric fields acting on drag-free test masses”. In: *Advances in Space Research* 39.2 (2007), pp. 213–218. ISSN: 0273-1177. DOI: https://doi.org/10.1016/j.asr.2006.03.045.

[26] M. Armano et al. “Charge-Induced Force Noise on Free-Falling Test Masses: Results from LISA Pathfinder”. In: *Phys. Rev. Lett.* 118 (17 Apr. 2017), p. 171101. DOI: 10.1103/PhysRevLett.118.171101.

[27] M. Armano et al. “Constraints on LISA Pathfinder’s self-gravity: design requirements, estimates and testing procedures”. In: *Classical and Quantum Gravity* 33.23 (Nov. 2016), p. 235015. DOI: 10.1088/0264-9381/33/23/235015.

[28] A. Cavalleri et al. “Gas damping force noise on a macroscopic test body in an infinite gas reservoir”. en. In: *Physics Letters A* 374.34 (July 2010), pp. 3365–3369. ISSN: 03759601. DOI: 10.1016/j.physleta.2010.06.041.

[29] Herbert B. Callen and Richard F. Greene. “On a Theorem of Irreversible Thermodynamics”. In: *Phys. Rev.* 86 (5 June 1952), pp. 702–710. DOI: 10.1103/PhysRev.86.702.

[30] M. Armano et al. “Charging of free-falling test masses in orbit due to cosmic rays: Results from LISA Pathfinder”. In: *Phys. Rev. D* 107 (6 Mar. 2023), p. 062007. DOI: 10.1103/PhysRevD.107.062007.

[31] P. J. Wass, T. J. Sumner, H. M. Araújo, and D. Hollington. “Simulating the charging of isolated free-falling masses from TeV to eV energies: Detailed comparison with LISA Pathfinder results”. In: *Phys. Rev. D* 107 (2 Jan. 2023), p. 022010. DOI: 10.1103/PhysRevD.107.022010.

[32] Simone Taioli et al. “The role of low-energy electrons in the charging process of LISA test masses”. In: *Classical and Quantum Gravity* (2023).

[33] M. Armano et al. “Sensor noise in LISA Pathfinder: An extensive in-flight review of the angular and longitudinal interferometric measurement system”. In: *Phys. Rev. D* 106 (8 Oct. 2022), p. 082001. DOI: 10.1103/PhysRevD.106.082001.

[34] G Hechenblaikner et al. “Digital Laser Frequency Control and Phase-Stabilization Loops in a High Precision Space-Borne Metrology System”. In: *IEEE Journal of Quantum Electronics* 47.5 (May 2011), pp. 651–660. ISSN: 0018-9197, 1558-1713. DOI: 10.1109/JQE.2011.2108637.

[35] Anna-Maria A. van Veggel and Christian J. Killow. “Hydroxide catalysis bonding for astronomical instruments”. In: *Advanced Optical Technologies* 3.3 (2014), pp. 293–307. DOI: doi:10.1515/aot-2014-0022.

[36] LISA Pathfinder collaboration. “Tilt-to-length coupling in LISA Pathfinder: a data analysis”. In preparation.

[37] Marie-Sophie Hartig, Sönke Schuster, and Gudrun Wanner. “Geometric tilt-to-length coupling in precision interferometry: mechanisms and analytical descriptions”. In: *Journal of Optics* 24.6 (June 2022), p. 065601. ISSN: 2040-8978, 2040-8986. DOI: 10.1088/2040-8986/ac675e.

[38] Marie-Sophie Hartig, Sönke Schuster, Gerhard Heinzel, and Gudrun Wanner. *Non-geometric tilt-to-length coupling in precision interferometry: mechanisms and analytical descriptions*. 2023.

[39] Marie-Sophie Hartig. “Tilt-To-Length coupling in LISA Pathfinder: model, data analysis and take-away messages for LISA”. PhD thesis. Gottfried Wilhelm Leibniz Universität Hannover, 2022. DOI: 10.15488/12113.

[40] M. Armano et al. “LISA Pathfinder platform stability and drag-free performance”. In: *Phys. Rev. D* 99 (8 Apr. 2019), p. 082001. DOI: 10.1103/PhysRevD.99.082001.

[41] J. I. Thorpe et al. “Micrometeoroid Events in LISA Pathfinder”. In: *The Astrophysical Journal* 883.1 (Sept. 2019), p. 53. DOI: 10.3847/1538-4357/ab3649.

[42] M. Armano et al. “LISA Pathfinder micronewton cold gas thrusters: In-flight characterization”. In: *Phys. Rev. D* 99 (12 June 2019), p. 122003. DOI: 10.1103/PhysRevD.99.122003.

[43] P. McNamara and G. Racca. *Introduction to LISA Pathfinder*. LISA-LPF-RP-0002. European Space Agency, 2009.

[44] M. Armano et al. “Temperature stability in the sub-milliHertz band with LISA Pathfinder”. In: *Monthly Notices of the Royal Astronomical Society* 486.3 (Apr. 2019), pp. 3368–3379. ISSN: 0035-8711. DOI: 10.1093/mnras/stz1017.

[45] LPF FCT. *LISA Pathfinder Report - May 2017*. Technical Note S2-ESC-RP-5664 Iss 1.0. ESA, July 2017.

[46] *LISA Technology Package Data Analysis, MATLAB Toolbox*.

[47] M. Armano et al. “Analysis of the accuracy of actuation electronics in the Laser Interferometer Space Antenna Pathfinder”. In: *Review of Scientific Instruments* 91.4 (2020), p. 045003. DOI: 10.1063/1.5140406.

[48] Gudrun Wanner and on behalf of the LISA Pathfinder collaboration Nikolaos Karnesis. “Preliminary results on the suppression of sensing cross-talk in LISA Pathfinder”. In: *Journal of Physics: Conference Series* 840 (May 2017), p. 012043. DOI: 10.1088/1742-6596/840/1/012043.

[49] Marie-Sophie Hartig and Gudrun Wanner. *Tilt-to-length coupling in LISA Pathfinder: analytical modelling*. 2023. DOI: 10.48550/arXiv.2305.03667.

[50] P. Sarra. *LISA Technology Package Core Assembly Gravitational Analysis Report*. Tech. rep. S2-CGS-RP-3001. Iss10. OHB-CGS, 2015.

[51] S. Vitale et al. “Data series subtraction with unknown and unmodeled background noise”. In: *Phys. Rev. D* 90 (4 2014), p. 042003. DOI: 10.1103/PhysRevD.90.042003.

[52] A. Papoulis. *Probability, random variables, and stochastic processes*. McGraw-Hill, 1991.

[53] Peter D. Welch. “The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms”. In: *IEEE Trans. Audio and Electroacoustics* 15.2 (1967), pp. 70–73.

[54] F.J. Harris. “On the use of windows for harmonic analysis with the discrete Fourier transform”. In: *Proceedings of the IEEE* 66.1 (1978), pp. 51–83. DOI: 10.1109/PROC.1978.10837.

[55] N. R. Goodman. “Statistical Analysis Based on a Certain Multivariate Complex Gaussian Distribution (An Introduction)”. en. In: *The Annals of Mathematical Statistics* 34.1 (Mar. 1963), pp. 152–177. ISSN: 0003-4851. DOI: 10.1214/aoms/1177704250.

[56] T. Ratnarajah and R. Vaillancourt. “Complex singular Wishart matrices and applications”. In: *Computers & Mathematics with Applications* 50.3 (2005), pp. 399–411. ISSN: 0898-1221. DOI: https://doi.org/10.1016/j.camwa.2005.04.009.

[57] G. Carter, C. Knapp, and A. Nuttall. “Estimation of the magnitude-squared coherence function via overlapped fast Fourier transform processing”. en. In: *IEEE Transactions on Audio and Electroacoustics* 21.4 (Aug. 1973), pp. 337–344. ISSN: 0018-9278. DOI: 10.1109/TAU.1973.1162496.

[58] Diane Valérie Ouellette. “Schur complements and statistics”. en. In: *Linear Algebra and its Applications* 36 (Mar. 1981), pp. 187–295. ISSN: 00243795. DOI: 10.1016/0024-3795(81)90232-9.

[59] C. R. Rao, ed. *Linear Statistical Inference and its Applications*. en. Wiley Series in Probability and Statistics. Hoboken, NJ, USA: John Wiley & Sons, Inc., 1973. ISBN: 9780470316436 9780471708230. DOI: 10.1002/9780470316436.

[60] Harold Jeffreys. “An invariant form for the prior probability in estimation problems”. In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 186.1007 (1946), pp. 453–461. DOI: 10.1098/rspa.1946.0056.

[61] Paul Shaman. “The inverted complex Wishart distribution and its application to spectral estimation”. en. In: *J. Multivar. Anal.* 10.1 (Mar. 1980), pp. 51–59.

[62] L. Svensson and M. Lundberg. “On posterior distributions for signals in Gaussian noise with unknown covariance matrix”. In: *IEEE Transactions on Signal Processing* 53.9 (2005), pp. 3554–3571. DOI: 10.1109/TSP.2005.853102.

[63] Michael Tröbs and Gerhard Heinzel. “Improved spectrum estimation from digitized time series on a logarithmic frequency axis”. In: *Measurement* 39 (Feb. 2006), pp. 120–129. DOI: 10.1016/j.measurement.2005.10.010.

[64] Justin Ellis and Rutger van Haasteren. *jellis18/PTMCMCSampler: Official Release*. Oct. 2017. DOI: 10.5281/zenodo.1037579.

[65] Joel N. Franklin. “Numerical Simulation of Stationary and Non-Stationary Gaussian Random Processes”. In: *SIAM Review* 7.1 (1965), pp. 68–80. ISSN: 00361445.

[66] Andrew Gelman et al. “Bayesian workflow”. In: *arXiv preprint arXiv:2011.01808* (2020).

[67] Paolo Chiggiato. *Outgassing properties of vacuum materials for particle accelerators*. https://arxiv.org/abs/2006.07124. 2020. DOI: 10.48550/ARXIV.2006.07124.

[68] P. A. Redhead. “Modeling the pump-down of a reversibly adsorbed phase. I. Monolayer and submonolayer initial coverage”. In: *Journal of Vacuum Science & Technology A* 13.2 (1995), pp. 467–475. DOI: 10.1116/1.579381.

[69] Karl Jousten, ed. *Handbook of vacuum technology*. eng. Second, completely revised and updated edition. Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA, 2016. ISBN: 9783527413386 9783527688241 9783527688234 9783527688265 9783527688258.

[70] Minxu Li and H. F. Dylla. “Model for the outgassing of water from metal surfaces”. en. In: *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films* 11.4 (July 1993), pp. 1702–1707. ISSN: 0734-2101, 1520-8559. DOI: 10.1116/1.578482.

[71] LISA Pathfinder collaboration. “Temperature-induced forces on geodesic reference masses for LISA: results from LISA Pathfinder”. In preparation.

[72] M. Armano et al. “Spacecraft and interplanetary contributions to the magnetic environment on-board LISA Pathfinder”. In: *Monthly Notices of the Royal Astronomical Society* 494.2 (Apr. 2020), pp. 3014–3027. ISSN: 0035-8711. DOI: 10.1093/mnras/staa830.

[73] S. Vitale. *Input to magnetic requirements definition for LISA*. LISA-UTN-INST-TN-002 Rev./Iss. 1.1. University of Trento, 2018.

[74] S. Vitale. *Revisitation of low frequency magnetic requirements for LISA*. UTN-LISA-INST-TN-015. University of Trento, 2020.

[75] Brigitte Kaune. “In-orbit stability analysis of the LISA pathfinder optical metrology: photoreceivers and polarisation”. PhD thesis. Hannover: Gottfried Wilhelm Leibniz Universität, 2021. DOI: 10.15488/10887.

[76] Q. Baghi et al. “Detection and characterization of instrumental transients in LISA Pathfinder and their projection to LISA”. In: *Phys. Rev. D* 105 (4 Feb. 2022), p. 042002. DOI: 10.1103/PhysRevD.105.042002.

[77] Hubert W. Lilliefors. “On the Kolmogorov-Smirnov Test for the Exponential Distribution with Mean Unknown”. In: *Journal of the American Statistical Association* 64.325 (1969), pp. 387–389. DOI: 10.1080/01621459.1969.10500983.

[78] L. Carbone et al. “Upper limits to surface-force disturbances on LISA proof masses and the possibility of observing galactic binaries”. In: *Physical Review D* 75.4 (2007).

[79] A Cavalleri et al. “A new torsion pendulum for testing the limits of free-fall for LISA test masses”. In: *Classical and Quantum Gravity* 26.9 (May 2009), p. 094017. ISSN: 0264-9381, 1361-6382. DOI: 10.1088/0264-9381/26/9/094017.

[80] L. Carbone et al. “Achieving Geodetic Motion for LISA Test Masses: Ground Testing Results”. en. In: *Physical Review Letters* 91.15 (Oct. 2003), p. 151101. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.91.151101.

[81] A Cavalleri et al. “Direct force measurements for testing the LISA Pathfinder gravitational reference sensor”. In: *Classical and Quantum Gravity* 26.9 (Apr. 2009), p. 094012. DOI: 10.1088/0264-9381/26/9/094012.

[82] Davide Dal Bosco. "Torsion Pendulum Testing of the LISA Charge Management System". PhD thesis. University of Trento, 2023. DOI: 10.15168/11572_375927.

[83] B P Abbott et al. "A guide to LIGO–Virgo detector noise and extraction of transient gravitational-wave signals". In: *Classical and Quantum Gravity* 37.5 (Feb. 2020), p. 055002. DOI: 10.1088/1361-6382/ab685e.

[84] C. Trenkel. *Gravity Analysis Report*. Tech. rep. S2.ASU.RP.2005. EADS-Astrium, 2015.

[85] B. Schläppi et al. "Influence of spacecraft outgassing on the exploration of tenuous atmospheres with in situ mass spectrometry". In: *Journal of Geophysical Research: Space Physics* 115.A12 (2010). DOI: https://doi.org/10.1029/2010JA015734.

[86] V. Ferroni and C. Zanoni. personal communication. Mar. 23, 2022.

[87] R. Kersevan and M. Ady. "Recent Developments of Monte-Carlo Codes Molflow+ and Synrad+". In: *Proc. 10th International Particle Accelerator Conference (IPAC'19), Melbourne, Australia, 19-24 May 2019* (Melbourne, Australia). International Particle Accelerator Conference 10. https://doi.org/10.18429/JACoW-IPAC2019-TUPMP037. Geneva, Switzerland: JACoW Publishing, June 2019, pp. 1327–1330. ISBN: 978-3-95450-208-0. DOI: doi:10.18429/JACoW-IPAC2019-TUPMP037.

[88] Marton Ady. "Monte Carlo simulations of ultra high vacuum and synchrotron radiation for particle accelerators". PhD thesis. École Polytechnique, Lausanne, 2016.

[89] Klaus Abich et al. "In-Orbit Performance of the GRACE Follow-on Laser Ranging Interferometer". In: *Phys. Rev. Lett.* 123 (3 July 2019), p. 031101. DOI: 10.1103/PhysRevLett.123.031101.

[90] H. D. Wealthy and Warren C. *Spacecraft-Level Gravity Model Results*. Tech. rep. S2.ASU.TN.2270.ISS02. EADS-Astrium, 2008.

[91] Dorothy M. Hoffman, Bawa Singh, and John H. Thomas, eds. *Handbook of vacuum science and technology*. San Diego, CA: Academic Press, 1998. ISBN: 9780123520654.

[92] C. D. Hoyle et al. "Submillimeter tests of the gravitational inverse-square law". In: *Phys. Rev. D* 70 (4 Aug. 2004), p. 042004. DOI: 10.1103/PhysRevD.70.042004.

[93] L. Carbone et al. "Thermal gradient-induced forces on geodesic reference masses for LISA". en. In: *Physical Review D* 76.10 (Nov. 2007), p. 102003. ISSN: 1550-7998, 1550-2368. DOI: 10.1103/PhysRevD.76.102003.

[94] LISA data challenge group. *LISA Data Challenge*. URL: https://lisa-ldc.lal.in2p3.fr/.

[95] B. P. Abbott et al. "GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral". In: *Phys. Rev. Lett.* 119 (16 Oct. 2017), p. 161101. DOI: 10.1103/PhysRevLett.119.161101.

[96] Henri Inchauspé, Martin Hewitson, Orion Sauter, and Peter Wass. "New LISA dynamics feedback control scheme: Common-mode isolation of test mass control and probes of test-mass acceleration". In: *Phys. Rev. D* 106 (2 July 2022), p. 022006. DOI: 10.1103/PhysRevD.106.022006.

[97] Daya K. Nagar and Arjun K. Gupta. "Expectations of Functions of Complex Wishart Matrix". en. In: *Acta Applicandae Mathematicae* 113.3 (Mar. 2011), pp. 265–288. ISSN: 0167-8019, 1572-9036. DOI: 10.1007/s10440-010-9599-x.

[98] Steven M Kay. *Fundamentals of statistical signal processing: Practical algorithm development*. Vol. 3. Pearson Education, 2013.

[99] A. Ortolan, G. Vedovato, M. Cerdonio, and S. Vitale. "Optimal reconstruction of the input signal in resonant gravitational wave detectors: Data processing algorithm and physical limitations". In: *Phys. Rev. D* 50 (8 Oct. 1994), pp. 4737–4743. DOI: 10.1103/PhysRevD.50.4737.

[100] C. Benvenuti. *Molecular surface pumping: the getter pumps*. Tech. rep. 43-50. CERN reports, 1999.