

Dealing with missing data: An inpainting application to the MICROSCOPE space mission.

Joel Bergé,^{1,*} Sandrine Pires,² Quentin Baghi,¹ Pierre Touboul,¹ and Gilles Métris³

¹ONERA - The French Aerospace Lab, 29 avenue de la Division Leclerc, 92320 Châtillon, France

²Laboratoire AIM, CEA/DSM-CNRS Université Paris Diderot IRFU/SAP,
CEA Saclay Orme des Merisiers 91191 Gif-sur-Yvette, France

³Geoazur (UMR 7329), Observatoire de la Côte d'Azur Bt 4,
250 rue Albert Einstein, Les Lucioles 1, Sophia Antipolis, 06560 Valbonne, France

(Dated: December 3, 2015)

Missing data are a common problem in experimental and observational physics. They can be caused by various sources, either an instrument's saturation, or a contamination from an external event, or a data loss. In particular, they can have a disastrous effect when one is seeking to characterize a colored-noise-dominated signal in Fourier space, since they create a spectral leakage that can artificially increase the noise. It is therefore important to either take them into account or to correct for them prior to e.g. a Least-Square fit of the signal to be characterized. In this paper, we present an application of the *inpainting* algorithm to mock MICROSCOPE data; *inpainting* is based on a sparsity assumption, and has already been used in various astrophysical contexts; MICROSCOPE is a French Space Agency mission, whose launch is expected in 2016, that aims to test the Weak Equivalence Principle down to the 10^{-15} level. We then explore the *inpainting* dependence on the number of gaps and the total fraction of missing values. We show that, in a worst-case scenario, after reconstructing missing values with *inpainting*, a Least-Square fit may allow us to significantly measure a 1.1×10^{-15} Equivalence Principle violation signal, which is sufficiently close to the MICROSCOPE requirements to implement *inpainting* in the official MICROSCOPE data processing and analysis pipeline. Together with the previously published KARMA method, *inpainting* will then allow us to independently characterize and cross-check an Equivalence Principle violation signal detection down to the 10^{-15} level.

PACS numbers: Valid PACS appear here: 07.05.Kf, 07.87.+v, 95.55.-n, 04.80.Cc

Keywords: Data analysis, Experimental test of gravitational theories

I. INTRODUCTION

Data acquisition in physics experiments is prone to intermittent cuts or errors. For instance, the observation strategy or commissioning periods, or maintenance works, may lead to gaps in the acquisition. Additionally, some events, either they are physical or instrumental, can lead to invalid data points that are *a posteriori* discarded, causing gaps in the data: examples of these kinds of gaps can be found in weak gravitational lensing, where cosmic rays and saturated pixels are masked (e.g. [1]), or in gravitational wave detection (e.g. [2]), where seismic events must be cut out.

Missing data complicate the data analysis by causing a spectral leakage in the frequency domain. This leakage is due to the convolution of the Fourier transform of the observational window with the Fourier transform of the complete data, causing the power from high-power regions to spread over the frequency domain. It can significantly hamper data analysis, as it has been shown in fields as different as Cosmic Microwave Background (CMB) analyses (e.g. [3–5]), weak gravitational lensing (e.g. [6–8]), asteroseismology (e.g. [9, 10]) or tests of fundamental physics (e.g. [11]). Optimizing the analysis

by taking into account irregular data sampling is particularly important in the case where small signal-to-noise (S/N) deterministic signals are to be detected and characterized, as will be the case e.g. for the MICROSCOPE (Micro-Satellite à traînée Compensée pour l'Observation du Principe d'Equivalence, [12–15]) or LISA Pathfinder [16, 17] missions. Efforts are underway to take missing data into account in data analysis, either by correcting them or designing data analysis methods insensitive to them (e.g. [18, 19] in weak lensing, [20] in CMB analyses, [10] in asteroseismology, [11] for MICROSCOPE). Among those efforts, *inpainting*, a method based on sparsity [21] has been developed in the last ten years, with a special emphasis first on 2D cosmology [3, 6], then on 1D astronomy [9, 10]. Thanks to its versatility, it can in principle be applied to many data analysis cases.

MICROSCOPE aims to test the Equivalence Principle in space with a 10^{-15} precision (at the 1σ confidence limit). The science data will consist of time series of differential accelerations measured by onboard inertial sensors. The data analysis challenge is to detect and characterize a possible very-small S/N Equivalence Principle Violation (EPV) signal. Despite its cutting-edge technology, it is expected that the data will be contaminated by uncontrollable events (such as impacts with micro-meteorites [11, 22]). We will have to discard the measurements corresponding to those events, therefore causing gaps in the usable time series, which we will

* joel.berge@onera.fr

need to take into account. The MICROSCOPE noise being strongly colored, it will be particularly subject to spectral leakage due to possible missing data when analyzing the data in Fourier space. A major part of the data analysis will consist in taking those missing values into account.

In this paper, we apply an *inpainting* reconstruction to mock MICROSCOPE data. In particular, we show how we are able to recover the Equivalence Principle violation signal after correcting missing data with *inpainting*, with no *a priori* assumption on the noise. Indeed, estimating an unknown colored noise is particularly difficult [23], especially when it is corrupted by missing data [11]. Therefore, we show that using a generic *inpainting* method, we are able to reach the MICROSCOPE's scientific goals.

This paper is organized as follows. Sect. II summarizes how missing data affect a general measurement and describes some advanced techniques developed to correctly deal with them; then Sect. III briefly presents *inpainting*. In Sect. IV, we summarize the MICROSCOPE mission. We apply *inpainting* to MICROSCOPE mock data in Sect. V; in particular, we show the gain in precision that we obtain in the evaluation of a possible Equivalence Principle violation signal with a Least Square fit when gaps are filled with *inpainting*. Sect. VI explores how *inpainting* depends on the observational window. We conclude in VII.

Unless otherwise noted, all quoted errors are 1σ errors.

II. MISSING DATA

A. Effect on the measurement: spectral leakage

Let $X(t)$ be the ideal regularly sampled complete time series, $Y(t)$ the measured incomplete time series and $M(t)$ the binary mask (with $M(t) = 1$ if we have information at data point $X(t)$, $M(t) = 0$ otherwise), such that $Y(t) = M(t)X(t)$. In the Fourier domain the multiplication of the mask becomes a convolution: the Fourier transform of the signal $\tilde{X}(f)$ is convolved by the Fourier transform of the mask $\tilde{M}(f)$,

$$\tilde{Y}(f) = \tilde{M}(f) * \tilde{X}(f), \quad (1)$$

At a given frequency, this convolution naturally produces a spectral leakage from this frequency to the frequencies around, the spectral leakage being characterized by the Fourier transform of the mask (the spectral window). The latter depending on the geometry of the mask, the spectral leakage also naturally depends on the mask geometry. For instance, a single hole produces a cardinal sine spectral (sinc) window, whose beam's width depends on the hole's width. For regularly distributed gaps, the Fourier transform of the mask is a Dirac comb, hence it introduces spurious peaks in the PSD [9, 10]. The case of a random mask is more complicated; as it is of interest for MICROSCOPE, we summarize it in the next subsection.

B. Randomly distributed missing values and colored noise

It can be shown that the spectral window for randomly distributed gaps of the same size, for a time series of length L , is asymptotically given by:

$$\tilde{M}(\omega) = \begin{cases} \left| \delta(\omega) - \frac{f_g L}{\sqrt{2\pi}} \right|, & (\omega = 0) \\ \frac{1}{2\sqrt{2}} \frac{f_g L}{\sqrt{N}} \left| \text{sinc} \frac{f_g L \omega}{2N} \right|, & (\omega \neq 0), \end{cases} \quad (2)$$

where N is the number of gaps, f_g the total fraction of missing values, ω the frequency, and $\delta(\omega)$ is the Dirac delta function.

Fig. 1 shows an example of spectral leakage due to randomly distributed gaps which could exist for MICROSCOPE (see Sect. IV for more details). The blue line represents the PSD of the noise in presence of gaps, obtained from the black curve by convolution with the mask (Eq. 2). The remaining lines are discussed in Sect. II C.

A short proof of Eq. (2) is given in the Appendix, together with a numerical example. From this equation, we can see that the spectral window corresponding to randomly distributed gaps of the same size has a complex dependence on the number of gaps and on the total masked fraction. In particular, far from its main peak ($\omega \rightarrow \infty$), the level of the spectral window depends only on the number of gaps N , and is independent of the total masked fraction f_g . Moreover, the loss of power in the main peak depends only on the total masked fraction. Numerical experiments show that this remains true when all gaps do not have the same size, in which case the sinc pattern disappears, but its envelope's level remains unchanged. Therefore, the spectral leakage far from its origin depends only on the number of gaps.

This is of particular importance for MICROSCOPE, since we will look for a deterministic signal, dominated by colored noise, far into the side lobes of the spectral window (as we describe in Sect. IV, the frequency range of the signal we seek is between 10^{-4} Hz and 10^{-3} Hz, while the spectral leakage originates from frequencies around 1 Hz). We will explore this behavior in Sect. VI.

C. Techniques to deal with missing data

A large number of studies have been presented in the literature that discuss various advanced solutions to the problem of missing data in time series. We summarize some of them below.

A widely used technique to estimate a spectral density on irregularly sampled data is the Lomb-Scargle periodogram [24, 25], which is based on a sine-wave least-square fitting. Assuming that a time series with missing data can be regarded as an irregularly sampled data set, we could view it as a way to correct for missing data. However, as mentioned in [10], this technique is subject to false detections, and we checked that it does not allow

us to correct for missing data in a noise-dominated time series.

The CLEAN algorithm [26, 27] is based on an iterative technique to detect peaks in the power spectrum, by detecting and removing them by descending amplitude in the Fourier domain, until the power spectrum is consistent with pure noise. However, any error on the estimated characteristics of removed peaks (amplitude, frequency, phase) introduces significant errors on the resulting final periodogram. As mentioned in [10], this is especially true when the spectral leakage is very large. It therefore cannot be used in our case of interest.

Recently, Baghi et al. [11] proposed the KARMA (Kalman-AR Model Analysis) method, a general linear regression method able to deal with incomplete data affected by unknown colored noise. It is based on an autoregressive (AR) fit of the noise that is used to whiten the data through a Kalman filtering process. By doing so the algorithm constructs a good approximation of the best linear unbiased estimator, conditionally to the AR model. The process is divided in three steps. First the AR parameters are estimated with Burg's algorithm [28] adapted to discontinuous data. The purpose of this procedure is to approximate an arbitrary noise spectral density with an AR model of suitable order. The second step is to build a vector of uncorrelated entries using the outputs of a Kalman filter. This avoids storing and inverting large correlation matrices. In the third step a least-squares regression is performed on the whitened data, providing a result with a quasi-minimal variance. The process may be iterated a few times so that both the estimated noise spectrum and the estimated regression parameters reach convergence.

Finally, a possibility to significantly decrease the spectral leakage consists in smoothing the window function, hence making it continuous: replace the abrupt changes from 1 to 0 (and vice-versa) around the i th hole (from times t_i^{start} to t_i^{end}) by a cosine function parametrized by a parameter τ , such as

$$M(t) = \begin{cases} 1 & \text{if } t < t_i^{start} - \tau \text{ or } t > t_i^{end} + \tau \\ 0 & \text{if } t_i^{start} \leq t \leq t_i^{end} \\ \frac{1}{2} \left[1 + \cos \left(\frac{\pi(t - t_i^{start} + \tau)}{\tau} \right) \right] & \text{if } t_i^{start} - \tau \leq t < t_i^{start} \\ \frac{1}{2} \left[1 + \cos \left(\frac{\pi(t - t_i^{end} - \tau)}{\tau} \right) \right] & \text{if } t_i^{end} < t \leq t_i^{end} + \tau \end{cases} \quad (3)$$

As shown by Fig. 1, re-defining the window function with Eq. (3) allows us to reduce the spectral leakage by up to a factor 5, depending on the value of τ . However, when using this definition, the window is not binary anymore, thus preventing us from using *inpainting* (introduced in the next section): although reducing the spectral leakage, it cannot help reconstruct the missing information with *inpainting*. Moreover, Eq. (3) corresponds to some weighting of the data, that is less optimal for Least-Square estimation than KARMA. Although this definition may help recover a high S/N signal, it is insufficient for our purposes (as we show below, the deter-

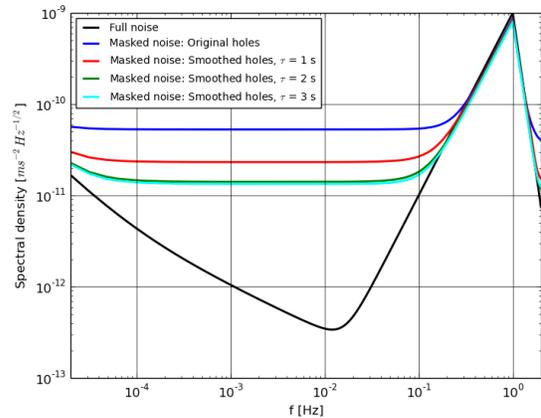


FIG. 1. Effect on the spectral leakage of smoothing the window function by defining it with Eq. (3). The blue line shows the spectral leakage with a binary mask, and other colored lines show the spectral leakage for smoothed windows, for varying τ .

ministic signal we seek remains below the most reduced spectral leakage).

III. INPAINTING

In this paper, we investigate a different approach to deal with missing data which is called *inpainting* in analogy with the restoration process used in museums to restore deteriorated paintings. Inpainting techniques are well known in the signal processing field and are currently used to fill gaps by inferring a maximum information from the remaining data. The classical inpainting problem can be defined as follows: having $Y = MX$, inpainting consists of recovering X knowing Y and M .

There is an infinite number of time series X that can perfectly fit the observed time series Y . We propose to use an inpainting method, introduced by [21], that relies on a prior of sparsity, which can be easily applied to MICROSCOPE data. This method makes use of the fact that all the representations of the signal are not equally interesting and pushes for sparse representations which makes information more concise and possibly more interpretable. This means that we have to find a representation $\alpha = \Phi^T X$ of the complete signal X (of length L) in the dictionary Φ^T (of size $L \times L$) where most coefficients α_i are close to zero, while only a few have a significant absolute value. For the MICROSCOPE application, Φ^T [29] can be the Discrete Cosine Transform (DCT) which is a decomposition into a set of oscillating functions like the Fourier transform. This DCT transform provides a sparse representation for the Equivalence Principle Violation signal and its calculation is efficient using Fast Cosine Transform algorithms that have the same computational complexity as Fast Fourier Transform algorithms.

Among all the possible solutions, we search for a unique solution by minimizing the number of large coefficient α in the selected dictionary Φ^T (i.e. that provides a sparse representation of the complete data) while imposing that the solution is equal to the observed data within the intrinsic noise of the data.

Thus, the solution is obtained by solving:

$$\min \|\alpha\|_1 \text{ subject to } \|Y - MX\|^2 \leq \sigma^2, \quad (4)$$

where $\|\cdot\|_1$ is the convex l_1 norm (i.e. $\|z\|_1 = \sum_k |z_k|$), $\|\cdot\|$ is the classical l_2 norm (i.e. $\|z\|^2 = \sum_k (z_k)^2$) and σ is the standard deviation of the noise in the observed time series. The solution of such a regularization can be approximated through an iterative algorithm called Morphological Component Analysis (MCA) introduced by [21]. This algorithm is based on a threshold that decreases exponentially at each iteration from a maximum value to zero. The conditions under which this algorithm provides an optimal and unique sparse solution to Eq. (4) have been explored by a number of authors (e.g. [30, 31]). They showed that the proposed method is able to recover the sparsest solution provided this solution is indeed sparse enough in the representation Φ^T and the mask is sufficiently random in this representation.

The *inpainting* technique has already been used to deal with missing data in several astrophysics applications (e.g. [6, 10, 32]).

IV. THE MICROSCOPE MISSION

Scheduled for launch in 2016, MICROSCOPE will test the Weak Equivalence Principle (WEP), a cornerstone of General Relativity, which states that two bodies in the same gravitational field experience the same acceleration, independently of their mass and composition. Thanks to a drag-free system and cutting-edge technology, MICROSCOPE will be able to test it down to the 10^{-15} level, 100 times better than the current best measurements [33], by looking for a possible Equivalence Principle Violation modulated signal in several long time series of noisy data.

A. Equivalence Principle Violation measurement principle

The principle of the measurement is to compare the acceleration experienced by two free-falling test masses in the Earth's gravity field. To this aim, MICROSCOPE embarks two ultrasensitive electrostatic differential accelerometers, each of these consisting of two coaxial cylindrical test masses whose motion is electrostatically constrained. The first accelerometer serves as a reference to demonstrate the experiment's accuracy: its test masses are made of the same material (Platinum-Rhodium PtRh10 alloy), so that it will not be subject to a true, physical EPV. The second accelerometer ('EP') is used to test the WEP: its test-masses are made of

different materials (PtRh10 and Titanium-Aluminium-Vanadium TA6V alloy). Any difference in the electrostatic servo-controlled acceleration applied to the masses to keep them relatively motionless is finely measured and provides the difference of the mass kinematic accelerations, a possible signal of the EP violation.

In orbit, the two masses of spherical inertia are precisely centered and they experience the same Earth gravitational field. Let us consider for simplicity a circular satellite orbit: in the instrument frame, the possible signal of violation, directed towards the Earth, is a sine whose frequency is defined by the orbital frequency, modulated by the spacecraft rotation.

The measured signal in the MICROSCOPE experiment can thus be written as

$$y_{EP}(t) = \frac{1}{2} \mathcal{M}_{EP} \delta g_{EP}(t) + \mathcal{S}(t) + \mathcal{N}(t), \quad (5)$$

where \mathcal{M}_{EP} is an instrumental calibration factor [13], δ is the EPV parameter we are aiming to detect and characterize, g_{EP} is the Earth gravity field's projection on the measurement-axis, \mathcal{S} represents systematics errors, and \mathcal{N} is the statistical inertial sensor noise.

The characteristics of the statistical noise \mathcal{N} can be predicted by performance analyses [13]. Its Power Spectral Density (PSD) can be decomposed in four regimes. A high-frequency increase as f^4 is due to the position detector noise; it is cut above 1 Hz by a Butterworth filter. The electronics noise is modeled as a 10^{-12} m/s²/Hz^{1/2} flat noise between 10^{-3} Hz and 0.03 Hz. At lower frequency, the noise is dominated by the gold wire used to keep test-masses at a constant electric potential, with a PSD variation as f^{-1} . Lower-frequency thermal fluctuations are modeled as a f^{-2} noise. The noise is assumed to be Gaussian, and centered on 0. On-ground performance measurements are in-line with those expectations. The black curve in Fig. 2 shows a simulation of the nominal noise as described above, smoothed with a 10-point wide Daniell filter. The peak around 2×10^{-4} Hz is a possible 10^{-15} EPV (to prevent it from being erased by the smoothing, the frequency region surrounding it is smoothed by a 5-point wide Daniell filter).

Nuisance and calibration parameters \mathcal{S} and \mathcal{M}_{EP} can be either corrected for through careful modeling (e.g. gravity gradient tensor) or estimated by in-flight calibration (e.g. bias [22, 34]), or minimized by design of the instrument and of the satellite (e.g. inertial tensor, instrument's noise).

Thus, assuming perfect correction of systematics and instrument's calibration, we will measure a colored-noise ($\mathcal{N}(t)$) dominated sine-wave signal ($\delta g_{EP}(t)/2$).

B. Missing values and spectral leakage

Different processes can cause data to be lost or not usable. A discussion about the sources of missing values in MICROSCOPE can be found in [11, 22]. They can be

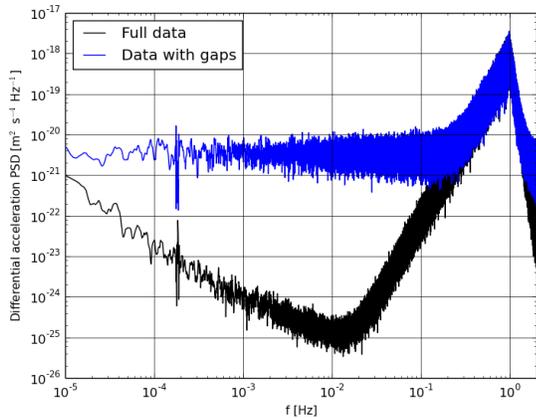


FIG. 2. Simulated MICROSCOPE’s nominal noise PSD, smoothed with a 10-point-width Daniell filter (black). An example of a possible EPV is shown by the peak at 2×10^{-4} Hz. The blue curve shows the PSD when gaps are present in the data.

divided in two main classes: (1) data transmission losses occur when a telemetry is not correctly received from the satellite; (2) saturated data occur when the instrument saturates (e.g. because of satellite’s multilayer insulation (MLI) coating crackles –as the temperature of the satellite’s sides varies according to their orientation towards or opposite the Earth along the satellite’s orbit–, or gas tank crackles –as the pressure of the tanks decreases as they are emptied by gas consumption–, or a micrometeorite impact). In the latter case, the data is not useable, and is therefore removed. In a worst case scenario, we expect that up to 3.5% of the data may be masked: most expected holes come from tank crackles ($\approx 90\%$), and MLI coating crackles ($\approx 10\%$), and are shorter than a second; micrometeorite impacts are expected to be rare and create very short gaps; teletransmission losses are wider (up to several seconds) but very rare.

The combined effect of those masks on the simulated noise PSD (Eq. 1) is shown by the blue line in Fig. 2: there is an obvious spectral leakage from $f \approx 1$ Hz to surrounding frequencies. As a result, the noise in the band $[10^{-4} - 10^{-1}]$ Hz is largely dominated by the spectral leakage from the high frequency noise: the deterministic signal we are looking for is therefore buried in the noise, while it would emerge clearly were all the data available.

Fig. 3 shows the spectral window $\tilde{M}(f)$ for each type of gaps described above. Their convolution with the signal’s PSD explains the leakage shown in Fig. 2. It is clear that the overall spectral window is dominated by that defined by tank crackles, while micrometeorite impacts and telemetry gaps have a negligible effect, their spectral window being two orders of magnitude smaller. As summarized in Sect. II, far from the main peak, the spectral window level depends only on the number of gaps, and is independent of the total masked fraction, in the case of randomly distributed gaps (as for MICRO-

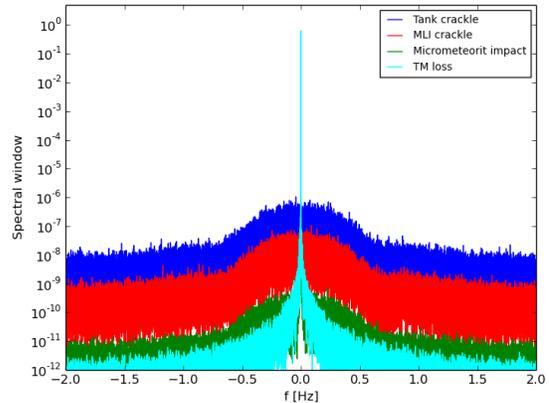


FIG. 3. Spectral windows corresponding to the possible gaps sources on MICROSCOPE.

SCOPE). This can be easily seen in Fig. 3. Hence, the level of the leakage shown in Fig. 2 depends only on the number of gaps (see also the Appendix).

V. INPAINTING APPLICATION TO MICROSCOPE

A. Simulations

To assess *inpainting*’s performance on MICROSCOPE-like data, we design a suite of simulations with the assumption that all nuisance parameters are perfectly corrected for: hence, the signal consists of just a pure sine at a well known frequency and noise with PSD as shown by the black lines of Fig. 2 and Fig. 4: $y_{EP}(t) = \delta g_{EP}(t)/2 + \mathcal{N}(t)$. For the sake of clarity in this section, and to have an acceptable S/N, we follow [11] and we set $\delta = 3 \times 10^{-15}$.

We simulate typical “inertial” measurement sessions consisting of 120 orbits, where the satellite’s attitude is kept fixed in an inertial frame. In this case, we expect an EPV signal at the orbital frequency, $f_{orb} = 1.8 \times 10^{-4}$ Hz, set by the mission design. Another type of measurement session consists in spinning the satellite about the axis normal to the orbital plane, so that the EPV signal is expected at a higher frequency (around $f_{EP} = 10^{-3}$ Hz): as the noise PSD is lower at this frequency (see Fig. 2), a 20-orbit integration is enough to detect a 10^{-15} EPV signal, were the data complete.

We define missing values in a worst case scenario, with 3% of missing values, due to 260 tank crackles per orbit, 24 MLI coating crackles per orbit, 0.2 micrometeorite impacts per orbit and 0.05 telemetry loss per orbit. We set the mean duration of saturated data due to crackles and micrometeorite impacts to 0.75 seconds (corresponding to 3 data points), and the telemetry losses can vary from 1 second to 250 seconds. Gaps are not pre-defined, but

their distribution is drawn randomly for each simulation, therefore we have access to their statistics only.

The black line in Fig. 4 shows the signal estimated PSD for a given simulation, when no mask is applied (i.e. the data is complete, regularly sampled), smoothed with a 20-point wide Daniell filter. Note that the smoothing decreases the apparent amplitude of the EPV peak. The inset on the left shows a zoom around f_{EP} , with no smoothing.

B. Inpainting

1. Missing data interpolation

The blue curve in Fig. 4 shows the measured signal in the Fourier domain when some data is missing. We see the same spectral leakage as previously mentioned. It is clear that the EPV signal is completely swamped in the spectral leakage, thereby becoming extremely difficult to recover without any correction when working in Fourier space.

The red curve shows the PSD measured after reconstructing missing values with *inpainting*. It is noticeable that the original PSD is not recovered correctly at intermediate frequencies. This may be the sign that MICROSCOPE data brings *inpainting* to its limits: although the expected EPV signal is indeed sparse, the noise may not be sparse enough for the *inpainting* reconstruction to be optimal. Despite this limitation, *inpainting* allows the EPV signal to emerge from the noise. It therefore becomes possible to detect and estimate the amplitude of this signal.

2. EPV signal detection and characterization

Once missing data have been reconstructed, we can use common techniques to detect and characterize the EPV signal: we indeed have a regularly sampled time series.

For instance, we can characterize the EPV signal with a regression analysis. Since, in this work, we assume that the signal we are looking for is a pure sine wave, of known frequency and phase, we perform a simple Least-Square fit in the Fourier domain to the corrected data to estimate its amplitude. For more realistic signals, including some errors in the nuisance parameters, as well as perturbing signal (e.g. a low frequency drift), a more general regression technique will be needed, like a Maximum-Likelihood method or a MCMC technique, which will allow us to constrain more parameters than just the EPV signal. Noise estimation is an upmost challenge in the Least-Square estimation of a signal (see e.g. [11, 23]). However, since in this work we are interested in evaluating the *inpainting*'s performance, we use a simple parametric fit of the noise, and do not look for the best possible noise estimator. We find that fitting the PSD of the noise as a sum of power laws (resembling the

MICROSCOPE specifications) provides us with a good enough noise estimation. Nevertheless, we should emphasize that given this sub-optimal noise estimator, the errors on the Least-Square estimators we give below are to be considered conservative.

For the case of interest of Fig. 4, where we simulated an EPV signal of 3×10^{-15} , we estimate $\delta = 3.51 \times 10^{-15} \pm 1.36 \times 10^{-15}$ after *inpainting* correction. In the case with no gaps, our Least-Square fit brings $\delta_{\text{complete}} = 3.89 \times 10^{-15} \pm 0.90 \times 10^{-15}$, while we measure $\delta_{\text{gap}} = 22.0 \times 10^{-15} \pm 17.1 \times 10^{-15}$ before the *inpainting* correction. This is a single example which is not enough to characterize the performances of the method.

Therefore, we then create 300 similar simulations to allow us to perform a statistical analysis of our estimates. In the remainder of this section, the errors we quote are the rms of the Least-Square estimators estimated on our 300 simulations; in this way, we are able to quantify the combination of errors coming both from the *inpainting* interpolation and from the Least-Square estimation. We find that after correcting for the missing values with *inpainting*, on average, we measure an EPV of $\delta = 2.69 \times 10^{-15} \pm 1.10 \times 10^{-15}$. With no *inpainting* correction, we would have measured $\delta_{\text{gap}} = 1.88 \times 10^{-15} \pm 13.5 \times 10^{-15}$ on average; in the case where no value is missing, $\delta_{\text{complete}} = 3.05 \times 10^{-15} \pm 0.76 \times 10^{-15}$. Therefore, *inpainting* allows us to have a significant (at better than 2σ confidence limit) measurement of a 3×10^{-15} EPV signal, which would be impossible by simply performing an ordinary Least Square fit on the available data. Given our estimated 1σ statistical error, we can conclude that we may be able to characterize a possible EPV signal with a 1.1×10^{-15} precision. Although slightly above MICROSCOPE requirements (10^{-15} precision), this is good enough for us to decide to implement *inpainting* in the official MICROSCOPE data analysis pipeline.

C. Comparison with KARMA

KARMA [11] and *inpainting* differ in their philosophy: *inpainting* fills gaps in a way that is independent of any physical model, thus allowing us to use an ordinary Least-Square method to estimate the EPV signal; KARMA, on the opposite, does not fill gaps but looks for an optimal estimator through a given model. Baghi et al [11] showed that the KARMA method applied to MICROSCOPE data allows us to measure a 3×10^{-15} EPV signal at the 99% confidence level in a similar worst case scenario than that used above. However, they assumed that the satellite was in “spinning” mode, where it spins about the axis normal to the orbital plane to increase the f_{EP} frequency, therefore moving it to a region where the instrumental noise is lower than that in the “inertial” mode that we investigated above.

In order to better compare KARMA and *inpainting*, we applied KARMA to the 300 simulations used above

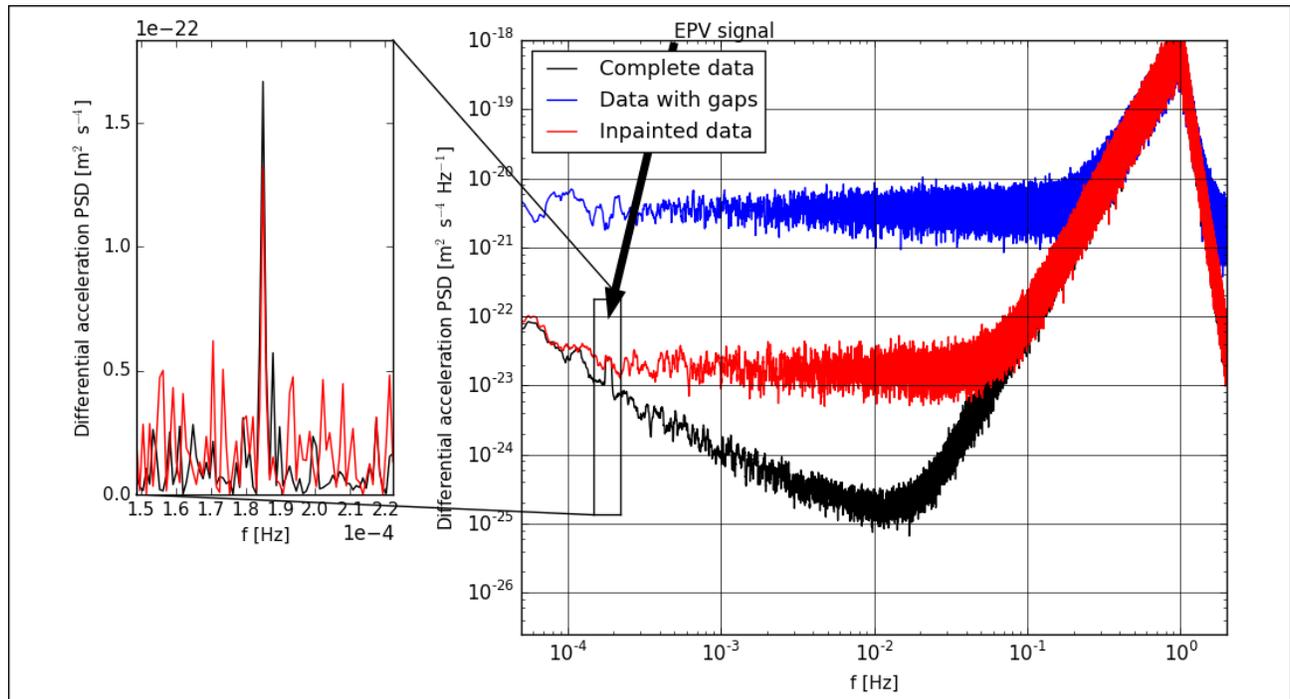


FIG. 4. MICROSCOPE differential acceleration PSD, smoothed with a 20-point Daniell filter, for a 120-orbit simulation. The black line shows the PSD when all the data is available, while the blue line shows the effect of missing values. The red curve shows the PSD after gaps are filled with inpainting. The inset on the left shows a zoom on the EPV signal, with no smoothing.

to test *inpainting*. In that scenario, we find that KARMA allows us to measure a 3×10^{-15} EPV signal as $\delta_{\text{KARMA}} = 3.06 \times 10^{-15} \pm 0.79 \times 10^{-15}$. Therefore, in the current “inertial” mode, KARMA’s performances are similar to what was shown in [11], and slightly better than those we estimated for *inpainting* above: on this mock data set, KARMA is more accurate and more precise than *inpainting*. However, as mentioned above, KARMA relies on an autoregressive model of the noise. This is a potential pitfall which might affect some data sets and which may be addressed in the future on additional data sets; in any case, it may require an independent data analysis to double check the results. *Inpainting*, although its performance is slightly poorer than KARMA’s, with its different philosophy, is an ideal tool for this purposes. It has thus been decided that the MICROSCOPE pipeline will rely both on KARMA and *inpainting* to analyze data and confirm a possible EPV detection.

Moreover, *inpainting* is much faster than KARMA: the full data analysis (missing values reconstruction and Least-Square estimation) on a typical 120-orbit session takes a few minutes on a Mac Pro (3.5 Ghz 6-core Intel Xeon E5 processor, 64 GB RAM), while the full KARMA processing on the same machine and data takes about 45 minutes. This speed brought us to implement *inpainting* as the primary missing value reconstruction tool for daily instrument monitoring when MICROSCOPE flies; *inpainting* being less precise than KARMA is not important in this monitoring, where only large signals will be

estimated, e.g. during calibration sessions [34].

VI. INPAINTING DEPENDENCE ON THE SPECTRAL WINDOW GEOMETRY: NUMBER OF GAPS AND MASKED FRACTION

In this section, we probe the *inpainting* dependence on the mask geometry when running on MICROSCOPE mock data; that is, we investigate how it allows us to detect and characterize a low-frequency deterministic signal, affected by a loss of power and a spectral leakage from high-frequency noise due to missing values.

To this aim, we design a set of simulations defining a grid in the (N, f_g) space, where N is the number of gaps per orbit and f_g the total fraction of missing values. As before, we simulate an 3×10^{-15} EPV signal as measured by MICROSCOPE, under the assumption that all nuisance parameters are perfectly corrected for. We set the (N, f_g) couples as to explore the parameter space around the nominal MICROSCOPE values (≈ 300 , $\approx 3\%$). Although they are overly unlikely for MICROSCOPE, we also investigate extreme values ($f_g \geq 30\%$) to increase the completeness of our exploration of *inpainting* performance. For each couple (N, f_g) , we create 100 simulations, run *inpainting* on them and measure the EPV signal with a Least-Square fit; our results are listed in Table I, where we report the mean measured EPV signal and statistical errors (multiplied by 10^{15}), defined as before, as the rms of the Least-Square estimators for each

set of 100 simulations. The couples with results noted “N/A” correspond to unphysical mask geometry (with a small masked fraction and a large number of gaps, which translates to gaps being smaller than one data point).

As mentioned in Sect. II (see also the Appendix), the spectral leakage far from the frequencies where the spectral leakage originates (at those frequencies where we look for the EPV signal in MICROSCOPE) depends only on the number of holes, and is independent of the total proportion of masked data. We then expect *inpainting* to have an easier job recovering the noise level in the case where the number of gaps is small, and therefore we expect our Least-Square estimate to be more precise. This is indeed what we observe in Table I: for a given total masked fraction f_g the error on the estimated EPV signal increases with the number of gaps. On the other hand, for a given number of gaps, the loss of power that affects the EPV signal itself because of missing values increases with the total masked fraction f_g (Eq. 2). Then, it will be more difficult for *inpainting* to reconstruct the EPV signal, and we expect the reconstruction to be biased. Indeed, we find that our Least-Square becomes less accurate as f_g increases.

From this study, we can draw some rough upper limit on the allowed number of gaps and total masked fraction to significantly measure a given EPV signal with MICROSCOPE, when correcting for missing values with *inpainting*: for example, with 300 gaps per orbit, the total masked fraction should not exceed $\approx 3\%$, which translates in gaps smaller than four data points (1 second) on average to characterize an EPV signal with a 1.08×10^{-15} precision. For a smaller number of gaps, a significant fraction can be lost (up to more than 60% with 30 gaps per orbit) while meeting the 10^{-15} MICROSCOPE requirement: this would imply large gaps (150 seconds) separated by large valid periods that provide uncontaminated information. Although such a gap configuration should not happen for MICROSCOPE, this number could be used to set the requirements of another, future mission.

This study brings us to design the following strategy to improve the missing values reconstruction: gather gaps with their immediate neighbors if they are separated by less than a given (ad hoc) distance, thus simplifying the spectral window’s geometry at the expense of removing valid data. Doing so, the number of gaps decreases, although the total masked fraction increases. To illustrate this, we create a set of 40 simulations, identical to those used above. The average linear power spectrum of the differential acceleration for complete and gaped data are shown in black and blue in Fig. 5, respectively. The red line shows the average linear power spectrum after the raw *inpainting* correction. We then increase the size of gaps by gathering those that are closer than 30 seconds: on average, the number of holes is decreased by a factor of 5, while the masked fraction goes from 3% to 50%. It is clear that as we decrease the number of gaps, the recovered noise after the *inpainting* correction

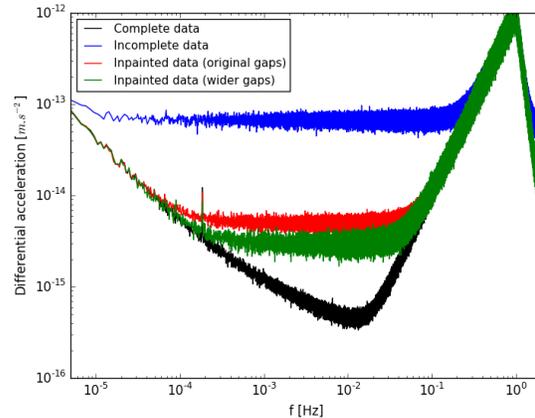


FIG. 5. Linear power spectrum of the mock MICROSCOPE differential acceleration, corrected with different masking strategies, averaged over 40 simulations. *Black*: complete data. *Blue*: data with missing values. *Red*: data corrected with *inpainting*, with no modification of the original gaps. *Green*: data corrected with *inpainting*, after gathering gaps closer than 30 seconds. Note the EPV peak at 1.8×10^{-4} Hz.

decreases (green line in Fig. 5), thereby improving the error on the Least-Square estimator; nevertheless, it is also clear from the figure that the amplitude of the recovered EPV signal peak decreases, thereby biasing the Least-Square estimator. Therefore, removing valid data to improve the signal characterization, although it allows us to improve the precision of our estimator, degrades its accuracy: a trade-off should then be found, that optimizes the precision and the accuracy of the Least-Square estimate. Characterizing it is beyond the scope of this paper, but may be of interest to optimize the MICROSCOPE data analysis and will be done in a further work.

VII. CONCLUSION

We have presented an *inpainting* application to MICROSCOPE data to correct for missing values. We first summarized the impact of missing data on a measured time series: in Fourier space, the observational window is convolved with the complete data, thereby creating a spectral leakage which artificially increases the noise at low frequency and hampers the detection and characterization of a small S/N deterministic signal. In particular, the use of an ordinary Least-Square estimation can be made impossible, as shown by [11]. Therefore, missing values need to be taken into account or corrected for in order to characterize a deterministic signal. This will be of utmost importance for the MICROSCOPE mission.

We then briefly mentioned some existing methods to deal with missing or irregularly sampled data; in particular, KARMA [11] has already been shown to perform well on MICROSCOPE data and reach the requirement to measure a 10^{-15} EPV signal with a 1σ confidence

TABLE I. EPV signal estimator and statistical error after *inpainting* correction, for different number of gaps N and total masked fraction f_g . Those numbers correspond to the mean and the rms (multiplied by 10^{15}) of the Least-Square estimators obtained on sets of 100 simulations for each couple (N, f_g) .

		N (per orbit)									
		30		100		300		1000		3000	
		$10^{15} \langle \delta \rangle$	$10^{15} \sigma_\delta$								
f_g	1.5%	3.10	0.83	3.12	0.88	2.82	1.06	N/A	N/A	N/A	N/A
	3%	2.96	0.83	2.79	0.86	2.89	1.08	N/A	N/A	N/A	N/A
	15%	3.06	0.82	2.94	0.90	2.57	1.44	5.43	3.36	3.93	2.28
	30%	2.84	0.88	2.50	1.13	2.69	1.73	6.67	4.55	8.12	5.71
	60%	2.40	1.08	1.80	1.17	2.77	2.21	9.40	10.2	23.4	22.1

limit. After shortly presenting *inpainting*, we applied it to mock MICROSCOPE data and showed that it allows us to correct for missing values in a worst-case scenario well enough to detect and characterize a 3×10^{-15} Equivalence Principle Violation signal at better than the 2σ confidence level. Our estimated statistical 1σ error of 1.1×10^{-15} is slightly bigger than required for the nominal MICROSCOPE performance (10^{-15}). This slight sub-performance certainly comes from the fact that although the expected EPV signal is indeed sparse in the DCT representation, the noise is not sparse enough, thereby limiting the *inpainting* reconstruction. The *inpainting* method reaches its limits with MICROSCOPE-type data. Further work to optimize *inpainting* in the MICROSCOPE case will be investigated in a future study. For instance, we could use a prior on the noise that could be inferred from a segment of the data without gaps. We are also currently investigating the feasibility of a hybrid technique between *inpainting* and KARMA, which is able to give an optimal estimator of the noise (Pires et al in prep).

We then explored how *inpainting* depends on the number of gaps and on the total masked fraction. This allowed us to set requirements on those two parameters for *inpainting* to reach MICROSCOPE's requirements on the EPV measurement. On the one hand, we saw that the precision of the Least-Square estimation of the EPV signal after an *inpainting* reconstruction mostly depends on the number of gaps; on the other hand its accuracy decreases with increasing total masked fraction, the estimator becoming biased low.

Finally, we showed that *inpainting's* performance is close to MICROSCOPE's requirements. As a consequence, we now have another tool to correct for missing values and characterize an EPV signal, that we can use in parallel with KARMA to reliably assert the detection and characterization of an Equivalence Principle violation signal. This study brought us to include *inpainting* in the official MICROSCOPE's data processing and analysis pipeline. Since it is fast and model-independent, it is well adapted for a quick missing data correction, needed for day-to-day instrument's monitoring and calibration. It will also be used along-side KARMA to estimate the EPV signal, allowing us to cross-check our results with two independent techniques.

ACKNOWLEDGMENTS

We thank Emilie Hardy, Manuel Rodrigues, Bruno Christophe and Alain Robert for useful discussions, as well as our anonymous referee for useful suggestions and comments. This work makes use of technical data from the CNES-ESA-ONERA-CNRS-OCA Microscope mission, and has received financial support from ONERA and CNES. We acknowledge the financial support of the UnivEarthS Labex program at Sorbonne Paris Cité (ANR-10-LABX-0023 and ANR-11-IDEX-0005-02).

Appendix: Spectral leakage, random gaps and colored noise

In this Appendix, we prove Eq. (2) and give a numerical example.

1. Analytical proof

We assume a time series of length $L \gg 1$, in arbitrary units. The window function that we aim to compute is defined in Sect. II: $M(t) = 1$ where data is available, $M(t) = 0$ where data is missing.

We further assume that data is missing in N non-overlapping holes, whose random distribution follows a uniform law. For simplicity, we assume that gaps are all of the same width Δ .

The mask function can be written:

$$M(t) = 1 - \sum_{k=1}^N \text{rect}_\Delta(t + n_k T), \quad (\text{A.1})$$

where n_k is a uniformly distributed random variable such as $n_{k+1} - n_k > \Delta > 1$ (non-overlapping holes), so that $n_{k+1}/n_k > 1$ and T is a given (ad hoc) time constant whose value is not important (since it is degenerate with the values of n_k).

Then, the spectral window is (with $\omega = 2\pi f$):

$$\tilde{M}(\omega) = \left| \delta(\omega) - \frac{\Delta}{\sqrt{2\pi}} \text{sinc}(\Delta\omega/2) \sum_{k=1}^N e^{-i\omega n_k T} \right|. \quad (\text{A.2})$$

In the case $\omega = 0$, the sum over k in Eq. (A.2) reduces to N and the spectral window is:

$$\tilde{M}(\omega)|_{\omega=0} = \left| \delta(\omega) - \frac{f_g L}{\sqrt{2\pi}} \right|, \quad (\text{A.3})$$

where we introduced the total masked fraction $f_g = N\Delta/L$. For a given time series, it depends on the total masked fraction only: it corresponds to the loss of power due to missing values at a given frequency.

In the case $\omega \neq 0$,

$$\tilde{M}(\omega)|_{\omega \neq 0} = \frac{\Delta}{\sqrt{2\pi}} \left| \text{sinc} \frac{\Delta\omega}{2} \right| \left| \sum_{k=1}^N e^{-i\omega n_k T} \right|. \quad (\text{A.4})$$

To estimate the trigonometric sum with random frequencies $\sum_{k=1}^N e^{-in_k T \omega}$ for each ω , we can define $\omega_k \equiv n_k T \omega$ and the complex variable $Z \equiv \sum_{k=1}^N e^{-i\omega_k}$ (as mentioned above, the value of T and ω do not impact the result, since n_k is a random variable). Then, according to [35], and remembering that $n_{k+1}/n_k > 1$ (so that $\omega_{k+1}/\omega_k > 1$), the real and imaginary parts of Z are approximately independent Gaussians with mean equal to zero and variance equal to $N/2$, for large N . Therefore, its modulus $|Z|$ follows a Rayleigh distribution, with mean $\langle |Z| \rangle = \sqrt{\frac{N}{2}} \sqrt{\frac{\pi}{2}}$ and variance $\text{Var}(|Z|) = \frac{N}{2} \frac{4-\pi}{2}$.

Therefore, we can approximate, for every ω , $\left| \sum_{k=1}^N e^{-in_k T \omega} \right| \approx \sqrt{\frac{\pi}{2}} \sqrt{\frac{N}{2}}$.

Then, the spectral window for irregularly spaced gaps on an infinite time series, is given by:

$$\tilde{M}(\omega)|_{\omega \neq 0} = \sqrt{N} \frac{\Delta}{2\sqrt{2}} \left| \text{sinc} \left(\frac{\Delta\omega}{2} \right) \right|. \quad (\text{A.5})$$

Returning to a finite time series of length $L \gg 1$, and substituting the total masked fraction f_g to Δ in Eq. (A.5), we finally obtain:

$$\tilde{M}(\omega)|_{\omega \neq 0} = \frac{1}{2\sqrt{2}} \frac{f_g L}{\sqrt{N}} \left| \text{sinc} \frac{f_g L \omega}{2N} \right|. \quad (\text{A.6})$$

Eqs. (A.3) and (A.6) then prove Eq. (2).

We now describe the behavior of the spectral window (Eq. 2) far from its main peak, in order to investigate how the total masked fraction f_g and the number of holes N affect its behavior. This regime is of interest in the case of MICROSCOPE.

Let us assume $\omega = \lambda\omega_s$, where $\omega_s = 1/L$ is the sampling frequency, with $\lambda = O(1)$. Then

$$\tilde{M}(\omega) = \frac{\sqrt{N}}{\lambda\omega_s\sqrt{2}} \sin \frac{f_g \lambda}{2N}. \quad (\text{A.7})$$

Therefore, the spectral window is bounded by $b = \frac{\sqrt{N}}{\lambda\omega_s\sqrt{2}}$, which does not depend on the total masked fraction, and depends only the total number of holes. At frequencies around the sampling frequency, the spectral window therefore scales as \sqrt{N} .

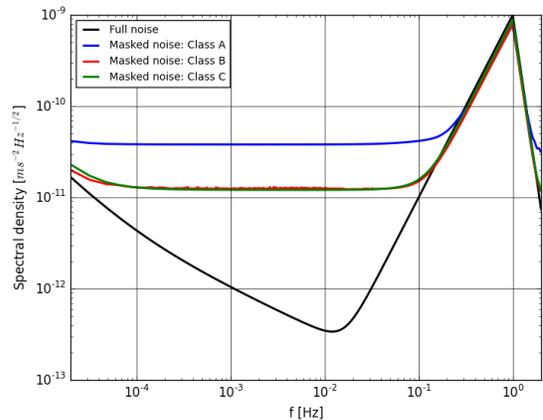


FIG. 6. Spectral leakage affecting the MICROSCOPE nominal differential acceleration's noise PSD (shown here is its square root) due to missing data. The black line shows the noise with no missing value, while the colored curves show the effect of missing data, depending on the geometry of the mask (classes A, B and C are shown in blue, green and red, respectively). The green curve completely overlaps the red one for frequencies less than 0.1 Hz.

The bound that we have just obtained, proportional to \sqrt{N} is true for every frequency (or equivalently, for every λ), becoming $b(\omega) = \frac{\sqrt{N}}{\omega\sqrt{2}}$.

2. Numerical example

To illustrate the behavior of the spectral window, we apply three classes of gaps to the expected MICROSCOPE differential acceleration noise. Those classes allow us to compare the effect of both the total fraction of masked data and the number of holes, and are defined as:

- Class A: 26,000 holes represent 20% of the total signal, with a size of 20 data points
- Class B: 2,600 holes represent 20% of the total signal, with a size of 200 data points
- Class C: 2,600 holes represent 2% of the total signal, with a size of 20 data points

The observed effect for each class of gaps is shown as the square root of the PSD in Fig. 6. The spectral leakage originates from the frequencies where the PSD is highest: mostly 1 Hz, but some spectral leakage from the lowest frequencies is also visible. Far from the frequencies from where the leakage occurs, we can observe that Classes B (red) and C (green) produce a similar spectral leakage, $\sqrt{10}$ lower than that of Class A (which has 10 times as many gaps than Classes B and C). It illustrates

that far from the origin of the leakage, the spectral leakage due to missing data depends only on the number of holes, as shown analytically above.

3. Discussion

We conclude this Appendix with a short summary of the behavior of the spectral window corresponding to randomly distributed gaps (Eq. 2) that go beyond our limited MICROSCOPE application:

- the spectral window at $\omega = 0$ depends on the total masked fraction f_g only; this corresponds to the loss of power due to gaps
- the behavior of the spectral leakage (linked to the spectral window at $\omega \neq 0$) at $\omega \ll 1$ (and $\omega \neq 0$) is governed by the interplay of the total masked fraction f_g and of the total number of gaps N
- the spectral leakage's envelope's level depends only on the total number of holes N for frequencies $\omega = O(1)$ and $\omega \gg 1$.

-
- [1] J. T. A. de Jong, G. A. Verdoes Kleijn, D. R. Boxhoorn, H. Buddelmeijer, M. Capaccioli, F. Getman, A. Grado, E. Helmich, Z. Huang, N. Irisarri, K. Kuijken, F. LaBarbera, J. P. McFarland, N. R. Napolitano, M. Radovich, G. Sikkema, E. A. Valentijn, K. G. Begeeman, M. Brescia, S. Cavuoti, A. Choi, O.-M. Cordes, G. Covone, M. Dall'Ora, H. Hildebrandt, G. Longo, R. Nakajima, M. Paolillo, E. Puddu, A. Rifatto, C. Tortora, E. van Uiter, A. Buddendiek, J. Harnois-Déraps, T. Erben, M. B. Eriksen, C. Heymans, H. Hoekstra, B. Joachimi, T. D. Kitching, D. Klaes, L. V. E. Koopmans, F. Köhlinger, N. Roy, C. Sifon, P. Schneider, W. J. Sutherland, M. Viola, and W.-J. Vriend, *ArXiv e-prints* (2015), arXiv:1507.00742.
- [2] M. Vallisneri, J. Kanner, R. Williams, A. Weinstein, and B. Stephens, *Journal of Physics Conference Series* **610**, 012021 (2015), arXiv:1410.4839 [gr-qc].
- [3] J.-L. Starck, D. L. Donoho, M. J. Fadili, and A. Rassat, *Astronomy & Astrophysics* **552**, A133 (2013), arXiv:1302.2758 [astro-ph.IM].
- [4] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. I. R. Alves, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, H. Aussel, and et al., *Astronomy & Astrophysics* **571**, A1 (2014), arXiv:1303.5062.
- [5] A. Rassat, J.-L. Starck, P. Paykari, F. Sureau, and J. Bobin, *JCAP* **8**, 006 (2014), arXiv:1405.1844.
- [6] S. Pires, J.-L. Starck, A. Amara, R. Teyssier, A. Réfrégier, and J. Fadili, *MNRAS* **395**, 1265 (2009), arXiv:0804.4068.
- [7] X. Liu, Q. Wang, C. Pan, and Z. Fan, *Astrophys. J.* **784**, 31 (2014), arXiv:1304.2873.
- [8] M. Shirasaki, N. Yoshida, and T. Hamana, *Astrophys. J.* **774**, 111 (2013), arXiv:1304.2164.
- [9] R. A. García, S. Mathur, S. Pires, C. Régulo, B. Bellamy, P. L. Pallé, J. Ballot, S. Barceló Forteza, P. G. Beck, T. R. Bedding, T. Ceillier, T. Roca Cortés, D. Salabert, and D. Stello, *Astronomy & Astrophysics* **568**, A10 (2014), arXiv:1405.5374 [astro-ph.SR].
- [10] S. Pires, S. Mathur, R. A. García, J. Ballot, D. Stello, and K. Sato, *Astronomy and Astrophysics* **574**, A18 (2015), arXiv:1410.6088 [astro-ph.SR].
- [11] Q. Baghi, G. Métris, J. Bergé, B. Christophe, P. Touboul, and M. Rodrigues, *Phys. Rev. D* **91**, 062003 (2015), arXiv:1503.01470 [gr-qc].
- [12] P. Touboul and M. Rodrigues, *Classical and Quantum Gravity* **18**, 2487 (2001).
- [13] P. Touboul, *Space Science Review* **148**, 455 (2009).
- [14] P. Touboul, G. Métris, V. Lebat, and A. Robert, *Classical and Quantum Gravity* **29**, 184010 (2012).
- [15] J. Bergé, P. Touboul, and M. Rodrigues, *Journal of Physics Conference Series* **610**, 012009 (2015), arXiv:1501.01644 [gr-qc].
- [16] P. McNamara, S. Vitale, K. Danzmann, and LISA Pathfinder Science Working Team, *Classical and Quantum Gravity* **25**, 114034 (2008).
- [17] M. Armano, M. Benedetti, J. Bogenstahl, D. Bortoluzzi, P. Bosetti, N. Brandt, A. Cavalleri, G. Ciani, I. Cristofolini, A. M. Cruise, K. Danzmann, I. Diepholz, G. Dixon, R. Dolesi, J. Fauste, L. Ferraioli, D. Fertin, W. Fichter, M. Freschi, A. García, C. García, A. Grynagier, F. Guzmán, E. Fitzsimons, G. Heinzl, M. Hewitson, D. Hollington, J. Hough, M. Hueller, D. Hoyland, O. Jennrich, B. Johlander, C. Killow, A. Lobo, D. Mance, I. Mateos, P. W. McNamara, A. Monsky, D. Nicolini, D. Nicolodi, M. Nofrarias, M. Perreur-Lloyd, E. Plagnol, G. D. Racca, J. Ramos-Castro, D. Robertson, J. Sanjuan, M. O. Schulte, D. N. A. Shaul, M. Smit, L. Stagnaro, F. Steier, T. J. Sumner, N. Tateo, D. Tomblato, G. Vischer, S. Vitale, G. Wanner, H. Ward, S. Waschke, V. Wand, P. Wass, W. J. Weber, T. Ziegler, and P. Zweifel, *Classical and Quantum Gravity* **26**, 094001 (2009).
- [18] C. Hikage, M. Takada, T. Hamana, and D. Spergel, *MNRAS* **412**, 65 (2011), arXiv:1004.3542 [astro-ph.CO].
- [19] J. T. VanderPlas, A. J. Connolly, B. Jain, and M. Jarvis, *Astrophys. J.* **744**, 180 (2012), arXiv:1109.5175.
- [20] L. Perotto, J. Bobin, S. Plaszczyński, J.-L. Starck, and A. Lavabre, *Astronomy & Astrophysics* **519**, A4 (2010).
- [21] M. Elad, J.-L. Starck, P. Querre, and D. Donoho, *Applied and Computational Harmonic Analysis* **19**, 340 (2005).
- [22] É. Hardy, A. Levy, G. Métris, M. Rodrigues, and P. Touboul, *Space Science Review* **180**, 177 (2013).
- [23] S. Vitale, G. Congedo, R. Dolesi, V. Ferroni, M. Hueller, D. Vetrugno, W. J. Weber, H. Audley, K. Danzmann, I. Diepholz, M. Hewitson, N. Korsakova, L. Ferraioli, F. Gibert, N. Karnesis, M. Nofrarias, H. Inchauspe, E. Plagnol, O. Jennrich, P. W. McNamara, M. Armano, J. I. Thorpe, and P. Wass, *Phys. Rev. D* **90**, 042003 (2014), arXiv:1404.4792 [gr-qc].

- [24] N. R. Lomb, *Astrophys. and Space Sci.* **39**, 447 (1976).
- [25] J. D. Scargle, *Astrophys. J.* **263**, 835 (1982).
- [26] D. H. Roberts, J. Lehar, and J. W. Dreher, *Astron. J.* **93**, 968 (1987).
- [27] G. Foster, *Astron. J.* **109**, 1889 (1995).
- [28] S. de Waele and P. M. T. Broersen, *IEEE Transactions on Signal Processing* **48**, 2876 (2000).
- [29] The transpose notation is the common notation in the literature of the domain for the direct representation.
- [30] M. Elad and A. Bruckstein, *IEEE Transactions on Information Theory* **48**, 2558 (2002).
- [31] D. Donoho and M. Elad, *Proc. Nat. Aca. Sci.* **100**, 2197 (2003).
- [32] E. Jullo, S. Pires, M. Jauzac, and J.-P. Kneib, *MNRAS* **437**, 3969 (2014), arXiv:1309.5718.
- [33] C. M. Will, *Living Reviews in Relativity* **17**, 4 (2014), arXiv:1403.7377 [gr-qc].
- [34] É. Hardy, A. Levy, M. Rodrigues, P. Touboul, and G. Métris, *Advances in Space Research* **52**, 1634 (2013).
- [35] R. Salem and A. Zygmund, *Proc. Nat. Acad. Sci.* **33**, 333 (1947).