Result of the MICROSCOPE weak equivalence principle test

Pierre Touboul, Gilles Métris, Manuel Rodrigues, Joel Bergé, Alain Robert, Quentin Baghi, Yves André, Judicael Bedouet, Damien Boulanger, Stefanie Bremer, et al.

▶ **To cite this version:**

Pierre Touboul, Gilles Métris, Manuel Rodrigues, Joel Bergé, Alain Robert, et al.. Result of the MICROSCOPE weak equivalence principle test. Classical and Quantum Gravity, 2022, 39 (20), pp.204009. 10.1088/1361-6382/ac84be . hal-03781319

HAL Id: hal-03781319
https://hal.science/hal-03781319

Submitted on 22 Sep 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Classical and Quantum Gravity

PAPER • OPEN ACCESS

Result of the MICROSCOPE weak equivalence principle test

To cite this article: Pierre Touboul *et al* 2022 *Class. Quantum Grav.* **39** 204009

View the article online for updates and enhancements.

You may also like

- Deflection and gravitational lensing of null and timelike signals in the Kiselev black hole spacetime in the weak field limit
Haotian Liu, Jinning Liang and Junji Jia
- The origin of Weyl gauging in metric-affine theories
Dario Sauro and Omar Zanusso
- On the relation between asymptotic charges, the failure of peeling and late-time tails
Dejan Gajic and Leonhard M A Kehrberger

IOP ebooks™

Bringing together innovative digital publishing with leading authors from the global scientific community.

Start exploring the collection–download the first chapter of every title for free.

This content was downloaded from IP address 88.127.95.218 on 22/09/2022 at 09:29

OPEN ACCESS
IOP Publishing
Classical and Quantum Gravity
Class. Quantum Grav. **39** (2022) 204009 (31pp)
https://doi.org/10.1088/1361-6382/ac84be

Result of the MICROSCOPE weak equivalence principle test

Pierre Touboul[1,†], Gilles Métris[2,*], Manuel Rodrigues[3,*], Joel Bergé[3], Alain Robert[4], Quentin Baghi[2,3,17], Yves André[4], Judicael Bedouet[5], Damien Boulanger[3], Stefanie Bremer[6,18], Patrice Carle[1], Ratana Chhun[3], Bruno Christophe[3], Valerio Cipolla[4], Thibault Damour[7], Pascale Danto[4], Louis Demange[2], Hansjoerg Dittus[8], Océane Dhuicque[3], Pierre Fayet[9,10], Bernard Foulon[3], Pierre-Yves Guidotti[4,19], Daniel Hagedorn[11], Emilie Hardy[3], Phuong-Anh Huynh[3], Patrick Kayser[3], Stephanie Lala[1], Claus Lämmerzahl[6], Vincent Lebat[3], Françoise Liorzou[3], Meike List[6,18], Frank Löffler[11], Isabelle Panet[12], Martin Pernot-Borràs[3], Laurent Perraud[4], Sandrine Pires[13], Benjamin Pouilloux[4,20], Pascal Prieur[4], Alexandre Rebray[3], Serge Reynaud[14], Benny Rievers[6], Hanns Selig[6,21], Laura Serron[2], Timothy Sumner[15], Nicolas Tanguy[3], Patrizia Torresi[4] and Pieter Visser[16]

[1] ONERA, Université Paris Saclay, F-91123 Palaiseau, France
[2] Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, IRD, Géoazur, 250 avenue Albert Einstein, F-06560 Valbonne, France
[3] ONERA, Université Paris Saclay, F-92322 Châtillon, France
[4] CNES Toulouse, 18 Avenue Edouard Belin—31401 Toulouse Cedex 9, France
[5] ONERA, Université de Toulouse, F-31055 Toulouse, France
[6] ZARM, Center of Applied Space Technology and Microgravity, University of Bremen, Am Fallturm, D-28359 Bremen, Germany
[7] IHES, Institut des Hautes Etudes Scientifiques, 35 Route de Chartres, 91440 Bures-sur-Yvette, France
[8] DLR, Köln headquarters, Linder Höhe, 51147 Köln, Germany

[*] Authors to whom any correspondence should be addressed.
[17] Current address: CEA, Centre de Saclay, IRFU/DPhP, 91191 Gif-sur-Yvette, France.
[18] Current address: DLR, Institute for Satellite Geodesy and Inertial Sensing, Am Fallturm 9, D-28359 Bremen, Germany.
[19] Current address: AIRBUS Defence and Space, F-31402 Toulouse, France.
[20] Current address: KINEIS, F-31520 Ramonville Saint-Agne, France.
[21] Current address: GERADTS GMBH, Kleiner Ort 8, D-28357 Bremen, Germany.
[†] Deceased in February 2021.

Original content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

[9] Laboratoire de physique de l'Ecole normale supérieure, ENS, Université PSL, CNRS, Sorbonne Université, Université Paris Cité, F-75005 Paris, France
[10] CPhT, Ecole Polytechnique, IPP, Palaiseau, France
[11] PTB, Physikalisch-Technische Bundesanstalt, Bundesallee 100, 38116 Braunschweig, Germany
[12] IPGP, 35 rue Hélène Brion, 75013 Paris, France
[13] Université Paris-Saclay, Université Paris Cité, CEA, CNRS, AIM Paris-Saclay, F-91190 Gif-sur-Yvette, France
[14] Laboratoire Kastler Brossel, Sorbonne Université, CNRS, ENS-PSL Université, Collège de France, 75252 Paris, France
[15] Blackett Laboratory, Imperial College London, Prince Consort Road, London, SW7 2AZ, United Kingdom
[16] Faculty of Aerospace Engineering, Delft University of Technology, Kluyverweg 1, 2629 HS Delft, The Netherlands

E-mail: gilles.metris@oca.eu and manuel.rodrigues@onera.fr

Received 28 January 2022, revised 22 July 2022
Accepted for publication 27 July 2022
Published 14 September 2022

Abstract

The space mission MICROSCOPE dedicated to the test of the equivalence principle (EP) operated from April 25, 2016 until the deactivation of the satellite on October 16, 2018. In this analysis we compare the free-fall accelerations ($a_{\rm A}$ and $a_{\rm B}$) of two test masses in terms of the Eötvös parameter $\eta({\rm A,B}) = 2\frac{a_{\rm A}-a_{\rm B}}{a_{\rm A}+a_{\rm B}}$. No EP violation has been detected for two test masses, made from platinum and titanium alloys, in a sequence of 19 segments lasting from 13 to 198 h down to the limit of the statistical error which is smaller than 10^{-14} for $\eta({\rm Ti,Pt})$. Accumulating data from all segments leads to $\eta({\rm Ti,Pt}) = [-1.5 \pm 2.3({\rm stat}) \pm 1.5({\rm syst})] \times 10^{-15}$ showing no EP violation at the level of 2.7×10^{-15} if we combine stochastic and systematic errors quadratically. This represents an improvement of almost two orders of magnitude with respect to the previous best such test performed by the Eöt-Wash group. The reliability of this limit has been verified by comparing the free falls of two test masses of the same composition (platinum) leading to a null Eötvös parameter with a statistical uncertainty of 1.1×10^{-15}.

Keywords: MICROSCOPE, general relativity, experimental gravitation, equivalence principle, space mission, space accelerometers, Eötvös parameter

(Some figures may appear in colour only in the online journal)

1. Introduction

The equivalence principle (EP) is the foundation stone on which Einstein built his new theory of gravitation, general relativity (GR) [1, 2]. GR has become an essential element in our description of the macroscopic Universe, from the big bang to black holes and gravitational waves. GR has passed with flying colours many stringent experimental tests (for reviews, see, e.g., [3] and chapter 21 in [4]).

However, fundamental physics is facing several conundrums which suggest the need to extend our present theoretical framework. On the gravity side, the missing mass problem

[5, 6], and the acceleration of the cosmic expansion [7, 8] have motivated the search for modifications of GR. On the particle-physics side, the peculiar structure of the Standard Model (SM), the hierarchy of particle masses, the observed preponderance of matter over antimatter, the presence of several different gauge symmetries with a curious symmetry breaking pattern, are some of the puzzles that motivate the search for extensions of the SM (notably supersymmetric ones [9]).

Most of the attempts to go beyond GR or beyond the SM, including the attempts to unify all interactions, have suggested the existence of new particles and of new interactions. In many cases, these new interactions give rise to apparent violations of the EP by predicting additional long-range feeble forces that do not couple, as Einsteinian gravity does, to the total mass–energy of a body. For instance, many theories including those with extra dimensions, from the Kaluza–Klein model [10, 11] up to string theories [12], suggest the existence of a light spin-0, dilaton-like, particle. Such a light scalar field can be made compatible with current Solar System tests if some screening mechanism is at work [13–21]. The coupling to matter of a dilaton-like particle is expected to violate the EP at a small level (see, e.g., references [14, 17, 22, 23]). Another possibility is the existence of a very light spin-1 U boson, related to an extension of the SM gauge group, mediating a new EP-violating force [24, 25].

The EP, or more precisely the weak equivalence principle (WEP) states that two bodies of different compositions and/or masses fall at the same rate in the same gravitational field (universality of free fall-UFF); equivalently, it states the equivalence of the 'inertial' and 'gravitational' masses. Since its use by Einstein in 1907 as a starting point of GR, it has been experimentally tested with higher and higher precision. Tests of the WEP are usually presented in terms of the Eötvös ratio η [26], defined as the normalised difference of accelerations (or equivalently, as the normalised difference of gravitational-to-inertial mass ratios) of two test bodies in same gravitational field [3]:

$$\eta(2,1) = 2\frac{a_2 - a_1}{a_2 + a_1} = 2\frac{m_{\mathrm{G}2}/m_{\mathrm{I}2} - m_{\mathrm{G}1}/m_{\mathrm{I}1}}{m_{\mathrm{G}2}/m_{\mathrm{I}2} + m_{\mathrm{G}1}/m_{\mathrm{I}1}} \tag{1}$$

where a_j is the acceleration of the jth test-body, and $m_{\mathrm{G}j}$ and $m_{\mathrm{I}j}$ are its gravitational and inertial masses. Since previous experiments have shown that $m_{\mathrm{G}}/m_{\mathrm{I}}$ does not differ from 1 by more than about 10^{-13}, the quantity which we directly measure,

$$\delta(2,1) = \frac{m_{\mathrm{G}2}}{m_{\mathrm{I}2}} - \frac{m_{\mathrm{G}1}}{m_{\mathrm{I}1}}, \tag{2}$$

differs from the Eötvös parameter $\eta(2,1)$ only by terms of order $[\eta(2,1)]^2$, and may in practice be identified with it.

Tests of the UFF have a long history, starting with Galileo Galilei (1638) and Newton (1687), and continuing to the end of the 20th century after Fischbach *et al* [27] revived the interest in experimental searches for new, WEP-violating interactions. The state-of-the-art experiments have measured $|\eta| <$ a few 10^{-13} (see reference [3] for a historical account of tests of the WEP): (i) the Eöt-Wash group used a high-precision torsion pendulum in the Earth and Sun gravitational fields [28, 29], and (ii) Lunar laser ranging has been used to monitor the motions of the Moon and the Earth around the Sun [30, 31], leading to a slightly better accuracy; but in this case η tests combination of effects due to composition differences between Earth and Moon (related to the WEP), and effects due to the self-gravity of each body (related to the strong EP).

Concepts for an EP test in space were first developed in Stanford (STEP project) to cope with ground experiments limitations [32, 33]. MICROSCOPE was the first space experiment to test the WEP and hence also UFF. Test masses in orbit follow quasi-infinite and purer free

falls in a quieter environment free of seismic disturbances and anthropogenic electromagnetic perturbations. The satellite was launched into a low-Earth, 710 km sun-synchronous orbit by a Soyuz rocket from Kourou on April 25, 2016. It delivered data for more than two years.

The satellite carries the twin space accelerometers for gravitation experiment (T-SAGE) payload (figure 1). T-SAGE is composed of two sensor units (SUs) called sensor unit for reference (SUREF) and sensor unit for the equivalence principle test (SUEP). Each SU includes two inertial sensors (or accelerometers), each one controlling one test mass. The test masses are concentric, co-axial hollow cylinders. Choosing cylindrical shapes for the test masses allows (i) to nest them with their centres of mass at the same position, (ii) to approximate their tensor of inertia to that of a sphere with appropriate choice of dimensions[22], and (iii) to optimise the capacitive sensing along the cylinders' common axis [34]. SUREF's test-masses are made of the same material (PtRh10), while SUEP's are made of different material (PtRh10) for the inner mass, Ti alloy for the outer mass [35]). The PtRh10 platinum–rhodium alloy contains 90% by mass of Pt ($A = 195.1$, $Z = 78$) and 10% of Rh ($A = 102.9$, $Z = 45$). The isotopic composition of Pt has been measured by PTB on a sample of flight material [35]. SUEP's outer test-mass is made of 90% titanium ($A = 47.9$, $Z = 22$), 6% of aluminium ($A = 27.0$, $Z = 13$) and 4% of vanadium ($A = 50.9$, $Z = 23$). The choice of the materials is a trade-off between machining laboratory know-how and theoretical motivation [36]. Titanium and platinum differ mainly from the neutron excess over the atomic mass $(N - Z)/A$ and, to a smaller extent, in the nuclear-electrostatic-energy parameter $Z(Z - 1)/(N + Z)^{1/3}$. The instrument works by measuring the electrostatic force required to equilibrate all other 'natural' forces in order to keep the test masses motionless with respect to electrodes fixed to the satellite [34]. The measured electrostatic force divided by the known mass is commonly called 'measured acceleration' (this is the opposite of the acceleration which would be undergone by the test mass in the absence of electrostatic force) and this terminology will be used in this paper. The driving idea of the experiment is to compare the measured accelerations of the two test mass pairs within each sensor to verify if they have the same free-falls. Along the X axis, parallel to the cylinder axis (figure 1, right panel and figure 2, right panel), capacitance changes are caused by the variation of the overlap between the test mass and its surrounding electrodes fixed to the satellite. Along the Y and Z axes the capacitance changes through variation of the gap. This allows for better sensitivity, removes electrostatic instability, and gives complete linearity along X. This is why the X-axis will be used in our analyses.

The MICROSCOPE satellite was designed to provide an environment as stable as possible. It was finely controlled along its six degrees of freedom with a drag-free and attitude control system (DFACS) described in [37]. The DFACS allows several modes of operation: inertial pointing or spin mode with the choice between several rotation rates. In all cases the instrument's X-axis is kept parallel to the orbital plane (figure 2, left panel). In inertial pointing mode the axes of the spacecraft and instrument are maintained pointing in a fixed direction and hence the direction of the Earth gravity field projected onto the X-axis varies at the orbital frequency and hence $f_{\rm EP} = f_{\rm orb}$. In the spin mode the satellite is rotated about the instrument Y-axis, which is orthogonal to the orbital plane, at a frequency $f_{\rm spin}$, and in this case $f_{\rm EP} = f_{\rm orb} + f_{\rm spin}$. The measurements analysed in this paper were obtained with two different spin rates referred to as $f_{\rm spin_2}$ and $f_{\rm spin_3}$. The values of the frequencies of interest in this paper are given in table 1 according to [38]. The higher frequency, $f_{\rm spin_3}$, is a trade-off between minimisation of instrument noise, which would favour a higher frequency, and the capabilities of

[22] But unlike a sphere, the momenta of order >2 are not null.

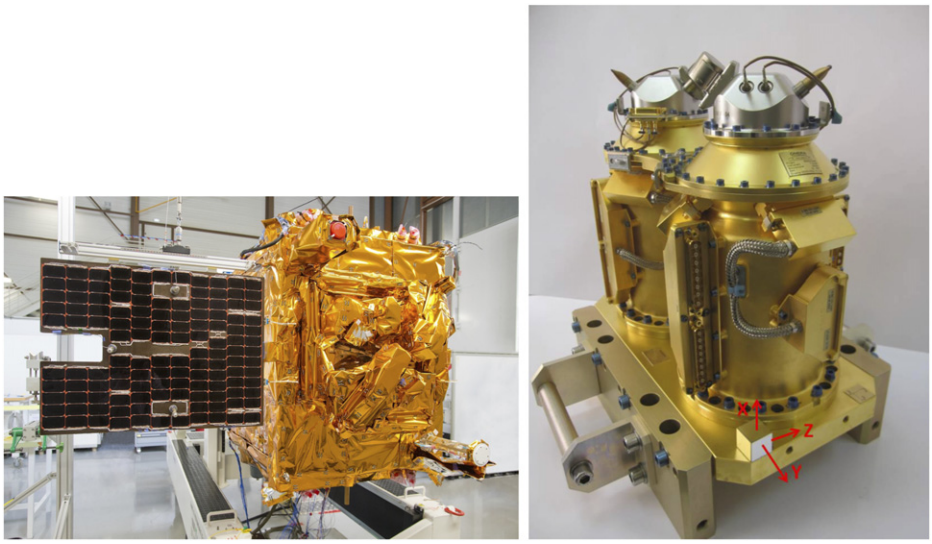

Figure 1. The MICROSCOPE satellite (left panel) and the T-SAGE instrument with its two cylindrical SUs (right panel).

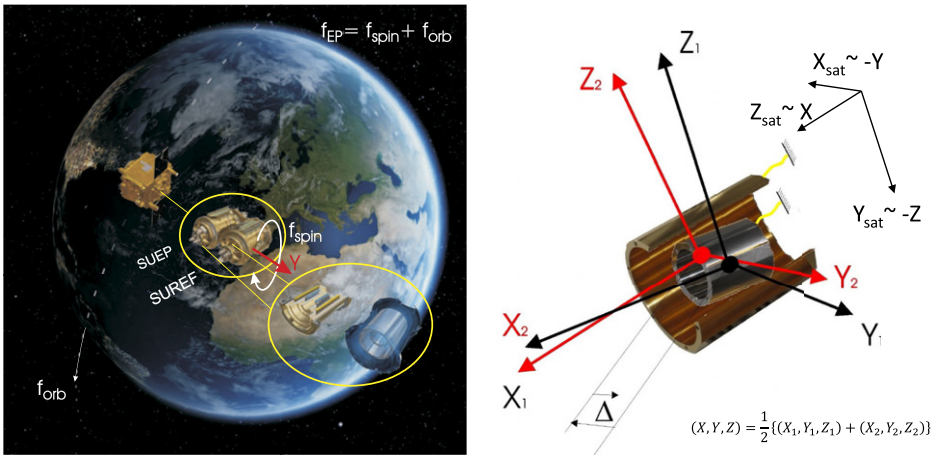

Figure 2. Left: the 4 test-masses orbiting around the Earth (credits CNES/Virtual-IT 2017). Right: reference frames for the satellite and for one pair of masses. The ($X_{\rm sat}$, $Y_{\rm sat}$, $Z_{\rm sat}$) triad defines the satellite frame; the reference frames (X_k, Y_k, Z_k, $k = 1, 2$) are attached to the test-masses (black for the inner mass $k = 1$, red for the outer mass $k = 2$); the X_k axes are the test-mass cylinders' longitudinal axis and define the direction of WEP test measurement; the Y_k axes are normal to the orbital plane, and define the rotation axis when the satellite spins; the Z_k axes complete the triads. The 7 μm gold wires connecting the test-masses to the common Invar sole plate are shown as yellow lines. $\vec{\Delta}$ represents the test-masses offcentring. The centers of mass have been approximately identified with the origins of the corresponding sensor-cage-attached reference frames.

Table 1. Main frequencies of interest.

Label	Frequency	Comment
$f_{\rm orb}$	$0.168\,18 \times 10^{-3}$ Hz	Mean orbital frequency
$f_{\rm spin_2}$	$\frac{9}{2} f_{\rm orb} = 0.756\,81 \times 10^{-3}$ Hz	Spin rate frequency 2 (V2 mode)
$f_{\rm spin_3}$	$\frac{35}{2} f_{\rm orb} = 2.943\,15 \times 10^{-3}$ Hz	Spin rate frequency 3 (V3 mode)
$f_{\rm EP_2}$	$0.924\,99 \times 10^{-3}$ Hz	EP frequency in V2 mode
$f_{\rm EP_3}$	$3.111\,33 \times 10^{-3}$ Hz	EP frequency in V3 mode

the micro-propulsion system; the smaller value, $f_{\rm spin_2}$, has the advantage of being farther from the limits of the propulsion system, and saves on gas usage.

In [39], we used 7% of the available data to provide first results. No evidence for a violation of the EP was found at the 1.3×10^{-14} level in terms of the Eötvös parameter, also providing improved constraints on additional new long-range forces [40–42]. Since then, the use of all data has allowed us to improve significantly the statistical error, and a thorough analysis of systematic errors has been conducted [43] using additional calibration sessions. The present paper describes these efforts and their result. In section 2, we review the available data at our disposal, grouped into scientific sessions. In section 3, we explain the physical parametric model used to analyse the measurements. In section 4 we show that perturbed behaviours of the measured accelerations lasting one or several orbits requires the division of some sessions into several disjointed segments. We list the characteristics of the segments analysed in this paper and describe the glitch (short singular events) detection strategy. The analysis is performed first on individual segments as shown in section 5, and then using the data from all segments gathered for a single estimation as presented in section 6. We conclude in section 7.

MICROSCOPE is simple in its principle but each component of the mission has been pushed to its limits given the external constraints (e.g. size of the satellite and global cost); for a more detailed presentation of MICROSCOPE the reader is referred to other papers of this volume (references [34, 37, 38, 43–48]).

2. Scientific sessions and available data

The MICROSCOPE observations are divided into different measurement sessions. A session represents a time span during which the satellite and the instrument keep the same configuration (spin, drag-free mode, etc). They are numbered by increasing integers and are described in the mission scenario [45]. Some of these sessions (called 'EP sessions') are directly devoted to the test of the EP while others ('calibration sessions') are used to calibrate or characterise the experimental apparatus [46]. EP sessions are the longest, most of them spanning 120 orbital periods (about 8 days), while calibration sessions typically span a few orbits. All sessions were performed with SUEP as well as with SUREF. Sessions are characterised by:

- The SU used: as explained in [34], the payload is composed of 2 SUs, each SU enclosing 2 co-axial concentric test masses. SUREF has 2 test masses with the same composition and is used as a null check of the experiment; SUEP aims at comparing the free falls of a test mass in platinum and of a test mass in titanium. During most of the sessions only one SU is on, with both operating simultaneously only during the EP sessions 430, 452 and 454.
- Which combination of the accelerometer outputs are used by the DFACS [37]: the DFACS uses micro-thrusters servo-controlled by the outputs of one or several inertial sensors in

Table 2. List of analysed sessions dedicated to the EP test with SUREF and their characteristics. Column 3 gives the calendar date of the beginning of each session, whereas columns 4 and 5 give the beginning and the end in terms of orbit number. The counting starts on 2016-04-26T01:03:05 UTC (orbit 1) and the orbital period is 5946 s. Columns 7 and 8 give the minimal and maximal temperatures measured in the SU during the session.

Session number	Spin	Beginning (date)	Beginning (orbit No.)	End (orbit No.)	Duration (orbits)	$T_{\min}$ (°C)	$T_{\max}$ (°C)
120	V2	2016-10-27	2681.5	2801.5	120.0	18.1	18.4
174	V2	2017-01-19	3952.8	4072.8	120.0	18.1	19.0
176	V2	2017-01-27	4074.4	4156.4	82.0	19.0	19.0
294	V3	2017-09-13	7397.6	7491.6	94.0	17.5	18.3
376	V2	2017-11-11	8258.0	8338.0	80.0	16.1	18.0
380	V3	2017-11-20	8382.6	8502.6	120.0	19.2	19.4
452	V2	2018-02-05	9507.7	9541.7	34.0	34.2	35.3
454	V2	2018-02-07	9542.6	9607.8	65.2	35.3	35.7
778	V2	2018-09-28	12 930.0	12 990.01	60.0	20.0	21.3

order to cancel their measured acceleration and to stabilise the rotation; it can be controlled by the output of one of the two inertial sensors (labelled IS1 for the inner mass and IS2 for the outer mass) within the SU in operation. A common mode use of both accelerometers is also possible. In practice almost all sessions used IS2 except sessions 358 and 406 which used the common mode. When both the SUs are working, the DFACS is controlled by the accelerometers within only one of the SUs: SUEP for session 430 and SUREF for sessions 452 and 454.

- The spin rate of the satellite: either V2 corresponding to the frequency $f_{\text{spin}_2} = (9/2)f_{\text{orb}} \simeq 7.57 \times 10^{-4}$ Hz or V3 corresponding to $f_{\text{spin}_3} = (35/2)f_{\text{orb}} \simeq 2.94 \times 10^{-3}$ Hz.
- The session duration: sessions were planned to last as long as possible, subject to operational constraints [45]: (i) periodic pointing updates required to remain in specification due to possible onboard clock drift; this limited the maximum duration to 120 orbits, each of duration $T_{\text{orb}} = 5946$; (ii) roughly once a month the star trackers pointed towards the bright moon and the fine attitude control had to be interrupted so that some sessions had to be shortened. In addition a few sessions were interrupted due to technical problems.

The first EP session after the commissioning is session 120, performed with SUREF. Note, the first calibration sessions revealed significant non-linearities in SUEP, which were solved by modifying the parameters of the proportional–integral–derivative controller in the servo-loop of SUEP, and its behaviour was nominal from session 210 onwards. None of the prior sessions are used in our analysis. Session 430 was interrupted by an anomaly and was discarded. Finally we are left with 9 EP sessions performed with SUREF (table 2) and 18 with SUEP (table 3).

Tables 2 and 3 list also the sensors' minimum and maximum temperatures. The typical temperature of the SUEP was about 10 °C while it was about 19 °C for the SUREF. The higher temperature in SUREF was due to two defective capacitors used for house-keeping data [34]. Note also the higher temperatures during sessions 452 and 454 when the two SUs ran simultaneously.

All sessions come with the following data used directly to estimate the Eötvös parameter [47]:

Table 3. Same as table 2 for sessions performed with SUEP.

Session number	Spin	Beginning (date)	Beginning (orbit No.)	End (orbit No.)	Duration (orbits)	$T_{\min}$ (°C)	$T_{\max}$ (°C)
210	V3	2017-02-14	4336.5	4386.5	50.0	9.7	10.4
212	V3	2017-02-18	4388.0	4464.1	76.1	10.4	10.4
218	V3	2017-02-28	4535.1	4655.1	120.0	9.1	10.3
234	V3	2017-03-15	4751.1	4843.1	92.0	9.3	10.3
236	V3	2017-03-21	4844.6	4964.6	120.0	10.3	10.4
238	V3	2017-03-29	4966.1	5086.1	120.0	10.4	10.6
252	V3	2017-04-13	5176.7	5282.7	106.0	9.2	10.7
254	V3	2017-04-20	5284.3	5404.3	120.0	10.7	10.8
256	V3	2017-04-29	5405.8	5525.8	120.0	10.8	10.8
326	V3	2017-09-27	7600.5	7702.5	102.0	10.1	12.9
358	V3	2017-10-14	7857.3	7950.1	92.8	9.4	9.9
402	V2	2017-12-06	8614.7	8634.7	20.0	10.0	10.3
404	V3	2017-12-07	8637.8	8757.8	120.0	10.3	11.7
406	V3	2017-12-16	8759.3	8779.3	20.0	11.7	11.7
438	V2	2018-01-16	9215.2	9255.2	40.0	9.8	10.7
442	V2	2018-01-22	9298.3	9338.3	40.0	11.2	11.4
748	V2	2018-09-03	12 562.3	12 586.3	24.0	9.3	9.7
750	V3	2018-09-05	12 589.3	12 597.3	8.0	9.7	9.8

- The measured accelerations (with the meaning explained in section 1) for each test mass of the operating SU at a sampling rate of 4 Hz with the associated time stamping. The difference of acceleration $\Gamma_x^{(d)}$ between the 2 test masses along the most sensitive axis X (see [34, 38] for the description of the axis), which was analysed in this work, is directly computed from these data.
- The attitude, angular velocity and angular acceleration of the satellite with respect to the inertial reference frame J2000 [37]. These data are given at exactly the same time stamps as the accelerometer measurements.
- The position and velocity of the centre of mass of the satellite in the J2000 frame sampled every minute.

Additionally, housekeeping data (sampled at 1 Hz) are used to monitor the behaviour of the experiment and to estimate systematic errors [43, 46]:

- The variation of position of each test-mass as measured by the capacitive sensors; the residual displacements are very small (less than 10^{-12} m at $f_{\rm EP}$) and the corresponding acceleration is negligible compared to our needs (section 5.5.4).
- The temperature which is measured by several probes inside the mechanical and electrical subsystems of each SU [34]. The corresponding systematic effects are estimated in [43].

3. The measurement model

A detailed explanation of the measurement model is given in [38]. Here we only summarise its main aspects. As explained in the introduction, we look for a difference of free fall between two co-axial concentric test masses of the same SU (SUREF or SUEP) by analysing the difference

of their measured accelerations $\overrightarrow{\Gamma}^{(d)} = \overrightarrow{\Gamma}^{(1)} - \overrightarrow{\Gamma}^{(2)}$. In a perfect experiment, this would correspond exactly to the applied differential acceleration $\overrightarrow{\gamma}^{(d)} = \overrightarrow{\gamma}^{(1)} - \overrightarrow{\gamma}^{(2)}$. Additional terms must be considered to account for the real experiment:

- A mapping matrix $\left[\tilde{\mathbf{A}}^{(c)}\right]$[23] between the applied and the measured acceleration; $\left[\tilde{\mathbf{A}}^{(c)}\right]$ is close to the identity matrix and takes into account scale factors and coupling between axes as well as a rotation reflecting the fact that our model for $\overrightarrow{\gamma}^{(d)}$ is not exactly expressed in the instrument frame which is imperfectly known: $\overrightarrow{\gamma}^{(d)} \rightarrow \left[\tilde{\mathbf{A}}^{(c)}\right]\overrightarrow{\gamma}^{(d)}$;
- A residual projection of the measured common mode acceleration $2\left[\tilde{\mathbf{A}}^{(d)}\right]\overrightarrow{\tilde{\Gamma}}^{(c)}$ where $\overrightarrow{\tilde{\Gamma}}^{(c)} = \left(\overrightarrow{\Gamma}^{(1)} + \overrightarrow{\Gamma}^{(2)}\right)/2$ and $||\left[\tilde{\mathbf{A}}^{(d)}\right]|| \ll 1$; formally, $\overrightarrow{\tilde{\Gamma}}^{(i)}$ are the noise-free measured accelerations but in practice, we approximate them with $\overrightarrow{\Gamma}^{(i)}$;
- A coupling with the angular acceleration $\dot{\overrightarrow{\Omega}}$ (because the sensors measure both linear and angular accelerations): $2\left[\mathbf{C}'^{(d)}\right]\dot{\overrightarrow{\Omega}}$;
- A measurement bias $\overrightarrow{B_0}^{(d)}$;
- A noise $\overrightarrow{n}^{(d)}$.

The origin of these terms is related both to detailed instrumental characteristics [34] and to the implementation of the instrument in its environment [38]. We end up with the model

$$\overrightarrow{\Gamma}^{(d)} = \overrightarrow{B_0}^{(d)} + \left[\tilde{\mathbf{A}}^{(c)}\right]\overrightarrow{\gamma}^{(d)} + 2\left[\tilde{\mathbf{A}}^{(d)}\right]\overrightarrow{\tilde{\Gamma}}^{(c)} + 2\left[\mathbf{C}'^{(d)}\right]\dot{\overrightarrow{\Omega}} + \overrightarrow{n}^{(d)}. \tag{3}$$

Note that this equation is fully similar to equation (19) of reference [38] but we have used the notation shortcut $\left[\tilde{\mathbf{A}}^{(d)}\right] = \left[\mathbf{A}^{(d)}\right]\left[\mathbf{A}^{(c)}\right]^{-1}$. The derivation of this model and in particular how we get a mixing between the applied differential acceleration and the measured common mode acceleration is detailed in [38]. Note also that other potential disturbing effects (non-linearity, thermal effects, stiffnesses and others) are not included in this model but are characterised in [43, 46] and contribute to the assessment of the systematic effects [43].

The difference of accelerations applied to the test masses derives directly from simple dynamics. First, each mass experiences Earth gravity; the gravitational force applied to each mass is slightly different because of (i) a gravity gradient due to their small difference of positions $\overrightarrow{\Delta}$ (called offcentring in the following, right panel of figure 2), and (ii) of a possible intrinsic difference of free fall parametrised by the Eötvös parameter. Second, since the acceleration is expressed in the instrument frame co-rotating with the satellite at an angular velocity matrix $[\Omega]$, the corresponding inertial acceleration must be taken into account. Third, additional small perturbations on the test masses (local gravity, magnetic effects, radiation pressure, radiometric effect and others) are gathered in the physical bias $\overrightarrow{b_1}^{(d)}$. The detailed derivation [38] yields

$$\overrightarrow{\gamma}^{(d)} = \delta(2,1)\overrightarrow{g}(O_{\rm sat}) + ([\mathrm{T}] - [\mathrm{In}])\overrightarrow{\Delta} + \overrightarrow{b_1}^{(d)}, \tag{4}$$

where

- $\delta(2,1) = m_{\mathrm{G}_2}/m_{\mathrm{I}_2} - m_{\mathrm{G}_1}/m_{\mathrm{I}_1} \simeq \eta(2,1) = -\eta(1,2)$ (note that equation (4) involves $\delta(2,1)$ instead of $\delta(1,2)$ because the measured differential acceleration is opposite to the difference of gravity accelerations),

[23] In the whole paper the notation [**M**] designates a matrix.

- $\overrightarrow{g}(O_{\rm sat})$ is the gravity acceleration computed at the centre of the satellite,
- [T] is the gravity gradient tensor computed at the centre of the satellite,
- $[\mathrm{In}] = [\Omega]^2 + [\dot{\Omega}]$, is the gradient of inertia matrix.

In this analysis, we will use only the measurement along the *X*-axis which is an order of magnitude more sensitive than those along *Y* and *Z*. This leads to

$$\begin{aligned}\Gamma_x^{(d)} \approx\ & B_{0x}^{(d)} + \tilde{a}_{c11} b_{1x}^{(d)} + \tilde{a}_{c12} b_{1y}^{(d)} + \tilde{a}_{c13} b_{1z}^{(d)} + \tilde{a}_{c11}\delta g_x + \tilde{a}_{c12}\delta g_y + \tilde{a}_{c13}\delta g_z \\ & + (T_{xx} - \mathrm{In}_{xx})\tilde{a}_{c11}\Delta_x + \big(T_{xy} - \mathrm{In}_{xy}\big)\tilde{a}_{c11}\Delta_y + (T_{xz} - \mathrm{In}_{xz})\tilde{a}_{c11}\Delta_z \\ & + \big(T_{yx} - \mathrm{In}_{yx}\big)\tilde{a}_{c12}\Delta_x + \big(T_{yy} - \mathrm{In}_{yy}\big)\tilde{a}_{c12}\Delta_y + \big(T_{yz} - \mathrm{In}_{yz}\big)\tilde{a}_{c12}\Delta_z \\ & + (T_{zx} - \mathrm{In}_{zx})\tilde{a}_{c13}\Delta_x + \big(T_{zy} - \mathrm{In}_{zy}\big)\tilde{a}_{c13}\Delta_y + (T_{zz} - \mathrm{In}_{zz})\tilde{a}_{c13}\Delta_z \\ & + 2\big(\tilde{a}_{d11}\Gamma_x^{(c)} + \tilde{a}_{d12}\Gamma_y^{(c)} + \tilde{a}_{d13}\Gamma_z^{(c)}\big) + 2\Big(c'_{d11}\dot{\Omega}_x + c'_{d12}\dot{\Omega}_y + c'_{d13}\dot{\Omega}_z\Big) \\ & + n_x^{(d)} - 2\big(\tilde{a}_{d11} n_x^{(c)} + \tilde{a}_{d12} n_y^{(c)} + \tilde{a}_{d13} n_z^{(c)}\big), \end{aligned} \tag{5}$$

where $\tilde{a}_{cij}$, $\tilde{a}_{dij}$, c'_{dij} are the elements of $\left[\tilde{\mathbf{A}}^{(c)}\right]$, $\left[\tilde{\mathbf{A}}^{(d)}\right]$ and $\left[\mathbf{C}'^{(d)}\right]$.

As explained in [47], some terms have been demonstrated to be negligible and others are corrected. Finally, after calibration and correction [47] (and see also section 5.1 for a summary), we get the fundamental equation which will be used for our analysis [47]:

$$\Gamma_{x,\mathrm{corr}}^{(d)} = \tilde{b}_x'^{(d)} + \delta_x g_x + \delta_z g_z + \Delta'_x S_{xx} + \Delta'_z S_{xz} + n_x^{(d)}, \tag{6}$$

where

- $\tilde{b}_x'^{(d)}$ is a bias which is almost constant but may slowly drift over time due to thermal effects;
- [**S**] is the symmetric part of the [**T**] − [**In**] matrix;
- $\delta_x = \tilde{a}_{c11}\delta \simeq \delta$ is very close to the Eötvös parameter since $|\tilde{a}_{c11} - 1| < 2 \times 10^{-2}$ whereas the potential contribution of the Eötvös parameter to $\delta_z = \tilde{a}_{c13}\delta$ should be much smaller because $|\tilde{a}_{c13}| < 2.6 \times 10^{-3}$ rad from manufacturing;
- Δ'_x and Δ'_z are effective combinations of the components of the offcentrings between the 2 test masses [47].

The structure of this equation is very simple:

- g_x, g_z, S_{xx} and S_{xz} are time-varying deterministic signals which can be computed accurately [47] knowing the position and the pointing of the satellite as well as its angular velocity and acceleration which are all delivered by CNES [38] with an accuracy better than the requirements [37];
- $\tilde{b}_x'^{(d)}$ is taken into account by estimating a polynomial trend;
- The parameters δ_x, δ_z, Δ'_x and Δ'_z are estimated.

These quantities are computed in the instrument frame, in which the varying signals of equation (4) have very different frequency patterns [36]:

- g_x and g_z are essentially periodic signals of frequency $f_{\rm EP}$ and are in phase quadrature,
- S_{xx} and S_{xz} have dominant components at DC and $2f_{\rm EP}$ and the variations of S_{xx} and S_{xz} at $2f_{\rm EP}$ are in phase quadrature,
- $\tilde{b}_x'^{(d)}$ is at very low frequency.

As a consequence, these signals are almost uncorrelated.

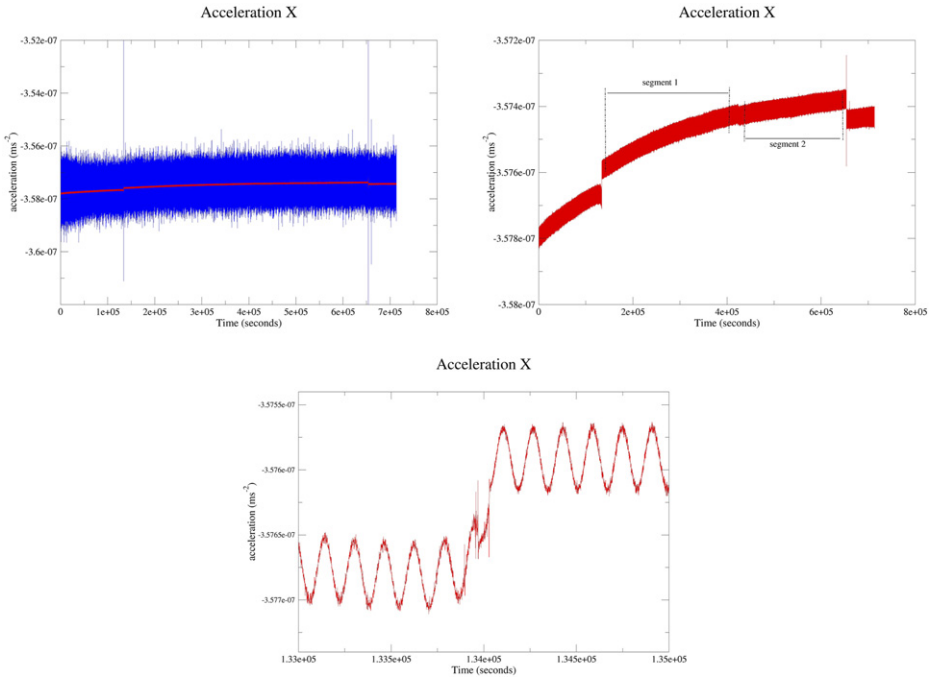

Figure 3. Differential acceleration for session 380. The first graph (top left) shows raw data in blue and filtered (using a running average over 60 s) data in red. The red curve reveals leaps which are not clearly visible on raw data: this appears more clearly by zooming in the ordinate axis as shown on the second graph (top right). Finally the third graph (bottom) which zooms in on the first leap shows that this is not a simple step but that there are disturbed measurements in the neighbourhood. Consequently, in the actual analysis, we use only data belonging to the two segments represented in the top right panel.

The bias, which encompasses the low frequency trend, is modelled with a degree 3 polynomial: $\tilde{b}'^{(d)}_x(t) = \sum_{j=0}^{3} \alpha_j (t - t_0)^j$. Substituting it in equation (6), we finally get

$$\Gamma^{(d)}_{x,\mathrm{corr}} = \sum_{j=0}^{3} \alpha_j (t - t_0)^j + \delta_x g_x + \delta_z g_z + \Delta'_x S_{xx} + \Delta'_z S_{xz} + n^{(d)}_x. \tag{7}$$

4. Handling of singular events

4.1. Segmentation

During some EP sessions, sudden changes in the local mean of the measured acceleration can be noted (figure 3). The typical macroscopic manifestation in the raw data (first panel of the figure) is a leap in the values of the measurements. These leaps are observed on the SUREF instrument only. They are not well understood as they are unpredictable, rare and not correlated to other observable events. Applying a lowpass filter and then zooming on the leap (third panel) shows that it does not consist of a simple Heaviside step function but is a complex mixing of erratic oscillations. We are faced with at least a few hundred seconds of unusable data. More

Table 4. Characteristics of the segments selected for our analysis of the SUREF data. The segment number corresponds to the session number extended by an index when there are more than one segment in the session. The duration is given as a multiple of orbital periods, remembering that this period is about 5946 s. The position of the segment in the session is indicated by the first and the last orbit which are included in the segment. The fourth column indicates the percentage of data eliminated from each segment during the pre-processing (see section 4.2.2).

Segment number	Duration (orbits)	Position in the session (orbits)	Percentage of data eliminated (glitches)
120-1	22	23 to 44	4
120-2	64	57 to 120	15
174	86	34 to 119	25
176	62	1 to 62	40
294	76	18 to 93	17
376-1	36	8 to 43	14
376-2	28	52 to 79	11
380-1	46	24 to 69	7
380-2	34	75 to 108	5
452	32	1 to 32	20
454	56	1 to 56	22
778-1	38	1 to 38	0
778-2	18	41 to 58	6

rarely some sessions have been stopped after the detection of technical problems (for example an instability in the drag-free loop) and there was a delay between the occurrence of the problem and the interruption. The data associated with the occurrence of such problems were discarded.

Thus, some sessions cannot be analysed in full, but only in parts referred to as 'segments' in the following. There can be a single segment if only the end of the session is corrupted or if the other potential useable parts are too short to bring a significant contribution; of course if no problem is detected in a session, the segment corresponds to the whole session. We can also extract several segments in one session as is the case for session 380 (figure 3). The driving principle is to end up with segments as long as possible including an even number of orbital periods: $T = 2nf_{\mathrm{orb}}$. Since f_{spin} has been chosen such that $f_{\mathrm{spin}} = (q/2)f_{\mathrm{orb}}$ (q odd integer), all potential signals at frequencies $f_i = k_i f_{\mathrm{orb}} + p_i f_{\mathrm{spin}}$ (with k_i and p_i any integers) combining the orbital frequency f_{orb} and the spin frequency f_{spin} are such that

$$f_i = k_i f_{\mathrm{orb}} + p_i f_{\mathrm{spin}} = \left(k_i + \frac{1}{2}p_i q\right) f_{\mathrm{orb}} = (2k_i + p_i q)n\frac{1}{T}. \tag{8}$$

This means that f_i corresponds to a sampling frequency of the discrete Fourier transform (DFT) and the correlation between two signals at frequencies f_i and f_j respecting the above property is null in theory and very low in practice. Tables 4 and 5 show the segments selected for our analysis. This comprises 13 segments totalling 598 orbits for SUREF and 19 segments totalling 1362 orbits for SUEP.

4.2. Glitches

4.2.1. General characteristics of glitches. When looking closely at the temporal evolution of the measured accelerations on several test-masses, we can see short (a few seconds) and

Table 5. Same as table 5 but for SUEP.

Segment number	Duration (orbits)	Position in the session (orbits)	Percentage of data eliminated (glitches)
210	50	1 to 50	18
212	60	1 to 60	17
218	120	1 to 120	15
234	92	1 to 92	18
236	120	1 to 120	21
238	120	1 to 120	24
252	106	1 to 106	26
254	120	1 to 120	27
256	120	1 to 120	28
326-1	66	2 to 67	12
326-2	34	69 to 102	7
358	92	1 to 92	14
402	18	3 to 20	35
404	120	1 to 120	23
406	20	1 to 20	23
438	32	1 to 32	21
442	40	1 to 40	21
748	24	1 to 24	25
750	8	1 to 8	19

significant (1–10 nm s^{-2}) variations which appear at the same time on both masses of the operating SU and even on the 4 masses when the 2 SUs are operating simultaneously (figure 4). The simultaneous appearance of these features for all masses proves that these events have the same common external source. This kind of event has been already observed in other space missions carrying accelerometers [49] and are called 'twangs' or 'glitches'. Glitches in MICROSCOPE have been extensively studied in a dedicated paper [48]. Here, we recall their main characteristics:

- As seen through the transfer functions of the drag-free and of the instrument, glitches look like exponentially damped sines with a large first ramp in one direction followed by a smoother oscillation in the opposite direction; their mean shape has been computed in [48] and is shown in figure 5; even if the source events probably do not last more than a few milliseconds, they affect the measured acceleration for a dozen seconds.
- The amplitude of the corresponding measured acceleration can reach up to a few 10^{-8} ms^{-2} but can be much smaller; there are probably also glitches which are masked by the measurement noise.
- The number of detected glitches typically ranges from 0.02 to 0.06 s^{-1}.
- Although they can occur at any time, their probability of occurrence is affected by two periodicities: the orbital period of the satellite and its spin period.

Their most likely origins are crackling of the MLI (multi-layer insulation) of the satellite and more rarely clangs of the gas tanks used for the micro-propulsion. Predicting the exact occurrence, form and amplitudes of the glitches seems out of reach. Moreover, although very similar, the responses of the different test masses to these sudden events are not perfectly identical and the corresponding signal is not fully cancelled in the differential acceleration [48].

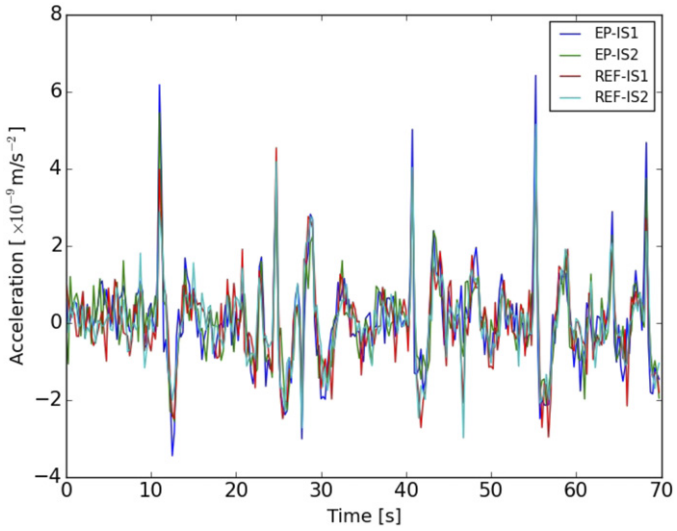

Figure 4. Example of superposition of measured acceleration (after removing of a low frequency trend) by the 4 test-masses at the same time. The large peaks appear on all masses but with slightly different amplitudes.

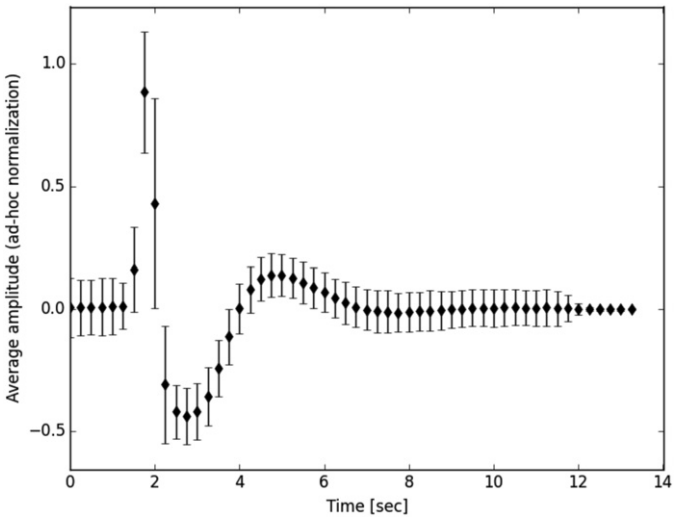

Figure 5. Mean shape of observed glitches [48].

Thus, due to the time distribution of the glitches, they could generate a tiny signal at the f_{EP} frequency. Consequently, the chosen strategy is to detect and eliminate them.

4.2.2. Detection and elimination of the glitches. Glitches are detected in a double two steps procedure [45, 47, 48]: (i) we use a standard recursive σ-clipping technique (e.g. reference [50]) to extract outliers (4.5σ) from the measured differential accelerations on the three axes (X, Y, Z) simultaneously before (ii) flagging data points in the second preceding the outliers and

the 15 s following it (typically, a single glitch builds from the noise within 0.5 s and dies off within 5–10 s from its peak, depending on its signal-to-noise ratio). Thence we define a first mask M_1 made of zeros in segments characterised as glitches and ones elsewhere. The same two steps are then performed on the high-frequency-filtered differential acceleration (using a 2nd order Butterworth filter of critical frequency 0.01 Hz), which allows us to detect low signal-to-noise glitches and define a second mask M_2 (a 3σ threshold is used during these steps). The final mask is the logical sum of both masks, $M = M_1 \times M_2$. The percentage of masked data for each session is indicated in tables 4 and 5. These are typically 20% but with less than 10% for some sessions and 40% for session 176.

5. Separated analysis of each segment

As explained in section 4.1 we have 13 segments for SUREF and 19 segments for SUEP. A first important objective is to analyse these segments separately since, as discussed in references [35, 39] the analysis of a single segment already resulted in an accuracy of estimation of the Eötvös parameter 10 times better than the previous experiments. Moreover this analysis will provide some insights into the data before a global analysis of all segments.

5.1. Main steps of the analysis

The analysis of individual segments is performed in the following steps:

(a) Calibration correction: thanks to dedicated calibration sessions, it was possible to estimate the values of the instrumental parameters Δ'_y, a_{d11}, a_{d12} and a_{d13} throughout the mission [43]; using the values of S_{ij} and $\dot{\Omega}_i$ precisely computed for each measurement date and of the common mode measured acceleration $\tilde{\Gamma}_i^{(c)}$, $\Gamma_x^{(d)}$ is corrected from the terms $\Delta'_y\left(S_{xy} + \dot{\Omega}_z\right)$ and $2\left(a_{d11}\tilde{\Gamma}_x^{(c)} + a_{d12}\tilde{\Gamma}_y^{(c)} + a_{d13}\tilde{\Gamma}_z^{(c)}\right)$ [47].

(b) Detection of the glitches: we detect glitches and define a corresponding mask according to the algorithm described in section 4.2.2. The union of this mask with the mask generated by the few points (typically a dozen per session) tagged directly on board is used during the analysis (see section 5.2) to discard the corresponding points. Note that this detection is performed only for EP sessions since we have checked that this operation has no impact on the parameters estimated using the calibration sessions.

(c) The parameters α_j, δ_x, δ_z, Δ'_x and Δ'_z are estimated by fitting the corrected measured differential acceleration to the model (7).

Since the measurement noise in MICROSCOPE is coloured (figure 6 and [35]), the optimal estimation of the parameters requires use of the characteristics of the noise. However, masking introduces gaps into the data, which are not regularly sampled any more. In this case, straightforward techniques (periodograms) to compute the power spectral density (PSD) are no longer effective: a significant leakage of the signal power from high to low frequency bands, leads to a substantial overestimation of the standard deviation of the estimated parameters [51, 52]. To solve this problem, we performed our estimation with two very different techniques described below: a method of maximisation of the likelihood in the time domain called modified expectation conditional maximisation (M-ECM) and a weighted least-squares regression in the Fourier domain named accelerometric data analysis for MICROSCOPE (ADAM).

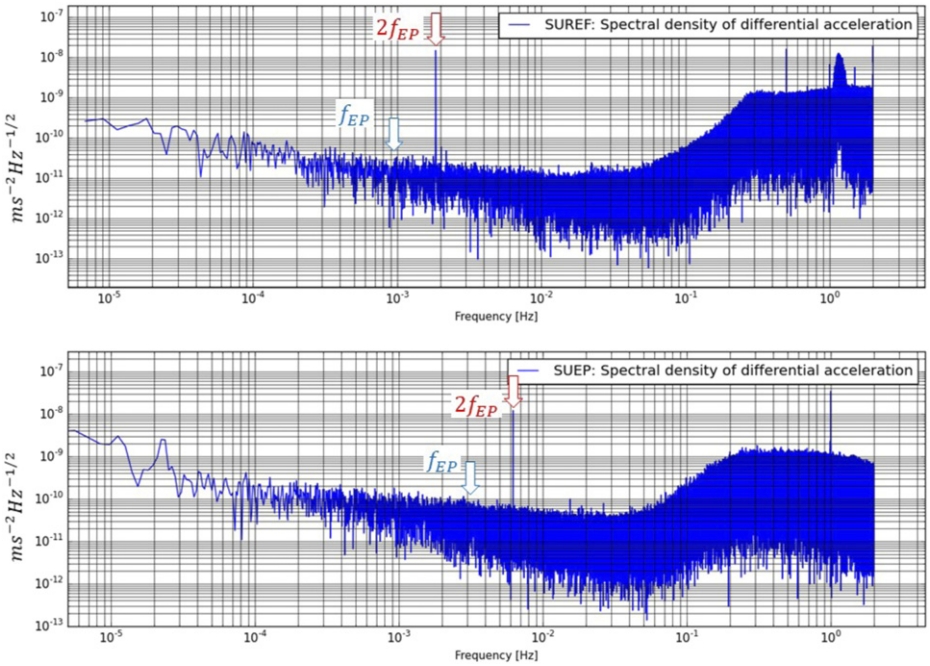

Figure 6. Amplitude spectral density of the sensor differential acceleration along the X-axis for SUREF during session 176 (upper panel) and for SUEP during session 236 (lower panel). Note the peak at frequency $2f_{\rm EP}$ coming from the Earth's gravity gradient due to the offcentring of the two test masses.

5.2. Estimation of the parameters in the time domain using the M-ECM analysis

M-ECM [52] is an inference algorithm designed to perform linear regression of gapped data. Although M-ECM may find applications in many areas, it was developed by the MICROSCOPE team during the mission design for the purpose of processing the mission data. Assuming a linear model for the signal and a stationary Gaussian distribution for the noise, it maximises the likelihood by iterating the following two steps. The first one is the conditional expectation step, which computes the expected likelihood (or rather, its logarithm $l_Y(\boldsymbol{\theta})$) conditionally to the observed data. This process amounts to estimating missing data within gaps. The second one is the maximisation step, which maximises the expected likelihood over the regression parameters $\boldsymbol{\theta}$. This step amounts to computing the generalised least-squares estimate of the parameters. It also includes estimating the noise PSD (the inclusion of this step motivates the qualification 'modified' to the algorithm's name, as it is not present in standard ECM algorithms [53, 54]). The use of the M-ECM algorithm allows one to avoid spectral leakage effects due to data gaps by restoring the statistics of the noise in the frequency domain.

In the case of complete data, and under the stationary assumption, we can write the logarithm of the likelihood as

$$l_Y(\boldsymbol{\theta}) \approx -\frac{1}{2}\left\{\log(\det[\boldsymbol{\gamma}]) + \left(\hat{\mathbf{Y}} - \left[\hat{\mathbf{A}}\right]\boldsymbol{\theta}\right)^{\dagger}[\boldsymbol{\gamma}]^{-1}\left(\hat{\mathbf{Y}} - \left[\hat{\mathbf{A}}\right]\boldsymbol{\theta}\right)\right\}, \tag{9}$$

where $\hat{\mathbf{Y}}$ is the vector of N Fourier-transformed measurements, $\boldsymbol{\theta}$ is the vector of parameters to estimate and $[\hat{\mathbf{A}}]$ is the matrix of derivatives of the model with respect to these parameters.

The matrix $[\gamma]$ denotes the covariance of the noise in the Fourier domain. It is approximately diagonal (see e.g. figure 1 of reference [47]) and its diagonal elements are equal to the noise PSD. If all measurements are available, we solve the estimation problem by maximising the likelihood with respect to $\boldsymbol{\theta}$.

When data points are missing, directly maximising equation (9) can be computationally cumbersome because we cannot approximate $[\gamma]$ as a diagonal matrix. Instead, the estimation is broken down in two steps. In the first step, we compute the expectation of the log-likelihood given the observed data $\mathbf{Y}_\mathrm{o}$ and the value of the parameters at the present iteration:

$$\text{E step: } Q_Y(\boldsymbol{\theta},\boldsymbol{\gamma}) = \mathrm{E}\left[l_Y(\boldsymbol{\theta})|\mathbf{Y}_\mathrm{o},\boldsymbol{\theta},\boldsymbol{\gamma}\right]. \tag{10}$$

This computation is called the expectation step (E). It requires computing the expectation of the full data vector conditionally on the observed data $\mathrm{E}[\hat{\mathbf{Y}}|\mathbf{Y}_\mathrm{o},\boldsymbol{\theta},\boldsymbol{\gamma}]$, along with its second-order moment $\mathrm{E}[\hat{\mathbf{Y}}\hat{\mathbf{Y}}^\dagger|\mathbf{Y}_\mathrm{o},\boldsymbol{\theta},\boldsymbol{\gamma}]$. We can compute these quantities by using the conditional mean and covariance formulas for Gaussian processes (see [52] for more details).

The second step is similar to the maximisation one would do for complete data, except that now we maximise the expected likelihood:

$$\text{M step: } \boldsymbol{\theta} = \underset{\theta}{\operatorname{argmax}}\, Q_Y(\boldsymbol{\theta},\boldsymbol{\gamma}). \tag{11}$$

This step is the maximisation (M) step. Note that solving for $\boldsymbol{\theta}$ is done for a given noise PSD $\boldsymbol{\gamma}$. In the M-ECM algorithm, we assume that $\boldsymbol{\gamma}$ is unknown and that it depends on some noise parameters $\boldsymbol{\beta}$. Hence, we also fit for the PSD by performing a pseudo maximisation conditionally on $\boldsymbol{\theta}$, so that

$$\boldsymbol{\beta} = \underset{\beta}{\operatorname{argmax}}\, Q_Y(\boldsymbol{\theta},\boldsymbol{\gamma}(\boldsymbol{\beta})). \tag{12}$$

Finally, we iterate E and M steps until $\boldsymbol{\theta}$ converges. The final solution is the value that maximises the likelihood with respect to observed data $l_{Y_\mathrm{o}}(\boldsymbol{\theta})$. As a result, M-ECM first computes the expectation of the likelihood through data reconstruction and then maximises it over the parameters. Hence, the maximisation takes advantage of the fast Fourier transform applied to the regularly sampled reconstructed time series, which is computationally more efficient than the direct maximisation of the gapped data likelihood. Values drawn from the missing data conditional distribution (also called reconstructed data in the following) is a useful by-product of M-ECM.

The resulting algorithm is unbiased after several iterations because it converges to the same solution as the one obtained with a direct (but costly) maximisation of the gapped, time-domain data likelihood. M-ECM is also approximately optimal in the statistical sense, as it yields the solution with nearly minimal variance, provided that the PSD estimation is sufficiently accurate [52].

5.3. Estimation of the parameters in the Fourier domain using the ADAM software

ADAM was developed by the MICROSCOPE team during the mission preparation for the purpose of processing the data of the experiment. For the estimation with ADAM, we use both original data remaining after the masking operation and data reconstructed by M-ECM as described in the previous section. This means that we have now data without gaps, i.e. regularly sampled. This allows us to apply a DFT converting the whole system of measurement equation in the frequency domain: whereas in the time domain each equation is associated to a time t_i,

Table 6. Values of the Eötvös parameter δ_x with their associated standard deviation estimated from SUREF measurements over individual segments. The table reports also the value of the component δ_z in phase quadrature. The values followed by (M) result from the M-ECM analysis while the values followed by (A) come from ADAM. Segments marked by an asterisk * correspond to sessions in spin V3 and the others to sessions in spin V2.

Segment number	δ_x (M) $\times 10^{15}$	δ_x (A) $\times 10^{15}$	δ_z (M) $\times 10^{15}$	δ_z (A) $\times 10^{15}$	δ_x/σ (M)	δ_x/σ (A)
120-1	-3.1 ± 16.7	-4.2 ± 15.9	13.5 ± 16.7	10.3 ± 15.9	-0.2	-0.3
120-2	-16.8 ± 8.5	-15.1 ± 8.3	-7.6 ± 8.5	-7.5 ± 8.3	-2.0	-1.8
174	7.8 ± 4.9	8.0 ± 4.2	-13.5 ± 4.9	-14.0 ± 4.2	1.6	1.9
176	1.7 ± 5.5	1.8 ± 4.5	7.9 ± 5.5	8.4 ± 4.5	0.3	0.4
294*	-8.0 ± 2.6	-7.7 ± 2.1	-2.8 ± 2.6	-2.3 ± 2.1	-3.1	-3.6
376-1	-3.4 ± 7.2	-4.1 ± 6.5	-8.5 ± 7.2	-7.9 ± 6.5	-0.5	-0.6
376-2	-5.7 ± 6.1	-6.4 ± 5.8	14.5 ± 6.1	16.1 ± 5.8	-0.9	-1.1
380-1*	7.6 ± 3.0	7.4 ± 2.4	-10.1 ± 3.0	-10.2 ± 2.4	2.5	3.1
380-2*	9.3 ± 3.1	8.9 ± 2.8	-9.1 ± 3.1	-9.2 ± 2.8	3.0	3.2
452	-4.3 ± 4.0	-4.8 ± 4.1	14.6 ± 4.0	16.1 ± 4.1	-1.1	-1.2
454	-3.1 ± 2.9	-3.7 ± 2.8	19.9 ± 2.9	19.8 ± 2.8	-1.1	-1.3
778-1	-8.1 ± 4.5	-8.1 ± 4.7	22.8 ± 4.5	22.6 ± 4.7	-1.8	-1.7
778-2	-2.3 ± 6.0	-3.2 ± 5.5	20.0 ± 6.0	18.6 ± 5.5	-0.4	-0.6

each equation of the transformed system is associated to a frequency f_j. This leads to several interesting properties and in particular:

- In case of periodic signals, their energy is concentrated in a small number of frequencies, corresponding to a small number of equations in the frequency domain; this is the case of the gravity acceleration and of the gravity gradient; moreover thanks to our choice to impose $f_{\rm spin} = kf_{\rm orb}/2$ and to perform the analysis over $2nT_{\rm orb}$ (k and n integers), the frequencies $f_{\rm orb}$, $f_{\rm spin}$ and $f_{\rm EP}$ correspond precisely to a frequency of the DFT: $f_q = q/T_{\rm analysis} = qf_{\rm orb}/(2n)$ with $q_{\rm orb} = 2n$, $q_{\rm spin} = kn$ and $q_{\rm EP} = (k+2)n$.
- With long enough data streams, the covariance is approximately diagonal, with a squared error inversely proportional to the data size [55]. In our application, the deviation is below the percent level. Thus a diagonal weighting matrix composed of the elements $w(f_k) = \frac{1}{\sqrt{\gamma(f_k)}}$ where $\gamma(f_k)$ is the PSD of the noise at the frequency f_k [47], is almost optimal.

Details of the procedure are described in [47] and we recall here the main steps:

(a) The series $\Gamma^{(d)}_{x,\rm corr}(t_i)$, $(t_i - t_0)^j$, $g_i(t_i)$, $g_z(t_i)$, $S_{xx}(t_i)$ and $S_{xz}(t_i)$ are transformed in the frequency domain by application of a DFT: the N observation equations in the time domain (corresponding to equation (7) at N different times) are transformed into N observation equations in the frequency domain.

(b) Potentially, we select only a subset of equations in the frequency domain, which is equivalent to selecting frequency bands. For the standard analysis, the frequency band around $f_{\rm EP}$ is selected to estimate δ_x and δ_z, and the frequency band around $2f_{\rm EP}$ is selected to estimate Δ'_x and Δ'_z. We choose a bandwidth large enough to encompass all the relevant signals (gravity acceleration and gravity gradient with their significant harmonics as well as the rotational terms): 8×10^{-4} Hz for sessions in spin V2 and 2×10^{-3} Hz for sessions in spin V3. The whole spectrum is used only to estimate the parameters of the low frequency trend described above.

Table 7. Same as table 6 but for SUEP.

Segment number	δ_x (M) $\times 10^{15}$	δ_x (A) $\times 10^{15}$	δ_z (M) $\times 10^{15}$	δ_z (A) $\times 10^{15}$	δ_x/σ (M)	δ_x/σ (A)
210*	-30.1 ± 14.5	-29.2 ± 13.1	-25.1 ± 14.5	-25.3 ± 13.1	-2.1	-2.2
212*	10.4 ± 13.9	9.5 ± 11.9	1.0 ± 13.9	1.3 ± 11.9	0.7	0.8
218*	3.6 ± 8.7	6.7 ± 8.1	8.2 ± 8.7	7.5 ± 8.1	0.4	0.8
234*	5.6 ± 9.3	5.9 ± 8.3	-4.2 ± 9.3	-4.6 ± 8.3	0.6	0.7
236*	2.7 ± 7.4	2.6 ± 6.6	16.7 ± 7.4	16.8 ± 6.6	0.4	0.4
238*	6.1 ± 7.8	5.8 ± 6.4	-3.9 ± 7.8	-3.7 ± 6.4	0.8	0.9
252*	-14.7 ± 8.7	-14.9 ± 7.3	1.8 ± 8.7	2.4 ± 7.3	-1.7	-2.0
254*	-14.2 ± 8.4	-14.1 ± 7.0	-24.9 ± 8.4	-25.8 ± 7.0	-1.7	-2.0
256*	-4.7 ± 8.6	-5.3 ± 7.4	14.1 ± 8.6	13.3 ± 7.4	-0.5	-0.7
326-1*	-10.1 ± 11.1	-16.3 ± 9.6	3.7 ± 11.1	2.3 ± 9.6	-0.9	-1.7
326-2*	-11.1 ± 15.4	-10.4 ± 13.5	5.5 ± 15.4	5.3 ± 13.5	-0.7	-0.8
358*	15.4 ± 11.9	15.8 ± 10.9	-2.0 ± 11.9	-1.9 ± 10.9	1.3	1.4
402	27.3 ± 35.1	28.4 ± 43.6	19.1 ± 35.1	28.7 ± 43.6	0.8	0.7
404*	6.3 ± 7.9	4.7 ± 6.7	-5.9 ± 7.9	-5.0 ± 6.7	0.8	0.7
406*	6.0 ± 18.6	5.9 ± 14.9	44.0 ± 18.6	44.1 ± 14.9	0.3	0.4
438	-12.5 ± 29.6	-23.4 ± 24.6	62.9 ± 29.6	54.5 ± 24.6	-0.4	-0.9
442	-10.7 ± 19.0	-1.5 ± 19.1	-5.2 ± 19.0	-6.3 ± 19.1	-0.6	-0.1
748	-17.5 ± 59.6	-23.4 ± 24.6	0.5 ± 59.6	54.5 ± 24.6	-0.3	-0.9
750*	66.6 ± 42.4	66.9 ± 38.4	23.6 ± 42.4	23.5 ± 38.4	1.6	1.7

(c) These observation equations are used to estimate the parameters of equation (7) and the associated statistical errors using a weighted least-squares inversion.

Fundamentally, no new information is expected from the ADAM analysis that performs the parameters estimation in the Fourier domain using the reconstructed data provided by M-ECM. However, it is interesting to cross-check results from different methods. Moreover, ADAM is much faster than M-ECM: for a segment of 120 orbits M-ECM needs about 12 h while ADAM takes only a few minutes. This is because in the E-step, M-ECM has to solve a large linear system where the system matrix is the covariance of the observed data in the time domain. The preconditioning is also memory-expensive. Although these disadvantages are not critical for a single run, they become a serious problem for the numerous tests required to strengthen the quality and the robustness of the estimation analysis.

5.4. Results

The estimates of the parameters δ_x and of δ_z and their standard deviations are listed in tables 6 and 7. As noted in section 3, $\delta_z = \tilde{a}_{c13}\delta$ is almost three orders of magnitude smaller than δ, i.e. far below what is detectable; thus the estimated values of δ_z and of its standard deviation are indicators of statistical or systematic effects. Figures 7 and 8 give an overview of the estimates of the Eötvös parameter with their time distribution. The values of the estimated offcentring are reported and commented in [43].

Several comments are in order.

(a) Although being very different algorithms, M-ECM and ADAM provide very similar results both for the values of δ_x and σ. The only appreciable difference is for δ_z estimated for segment 748 which is a SUEP session in spin V2 lasting 24 orbits; but this difference

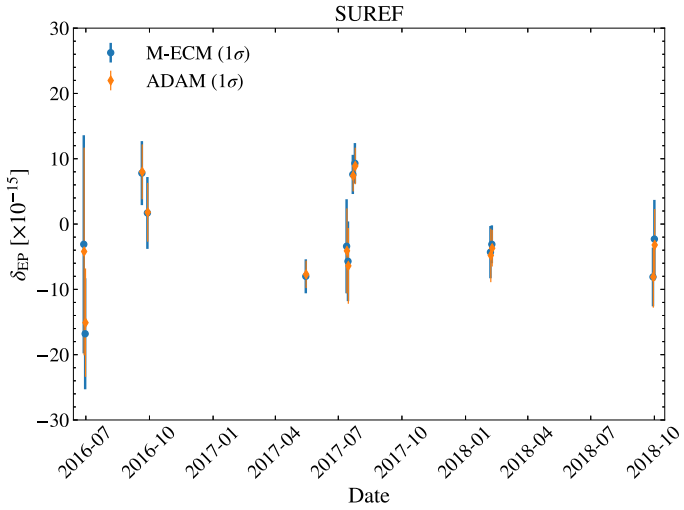

Figure 7. Eötvös parameter estimates for each SUREF segment and their 68% confidence error bars. Blue circles show M-ECM's estimates and orange ones ADAM's.

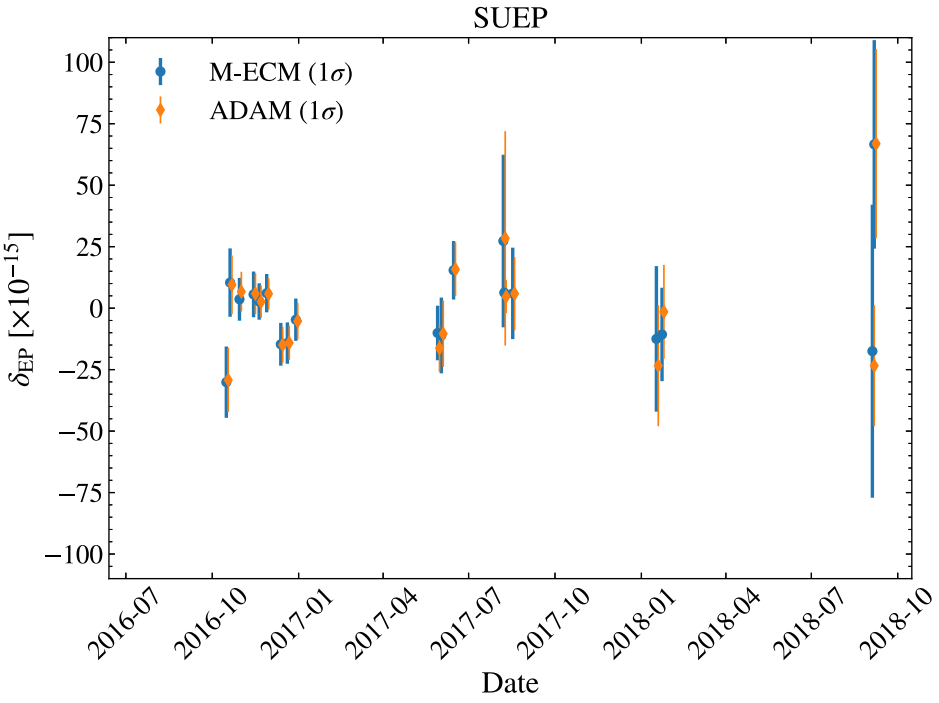

Figure 8. Same as figure 7 but for SUEP.

is statistically insignificant since it is smaller than the standard deviation computed by M-ECM.

(b) As expected given the frequency dependence of the noise visible on figure 6, the standard deviation is smaller in spin V3. Furthermore it logically depends on the duration of the analysis segment: it is shown in [46], that when normalised to the same duration (or equivalently, expressed in terms of PSD) all sessions with the same SU and the same spin have quite similar noise.

(c) As expected before the launch, SUREF provided more accurate data than SUEP, mainly thanks to its heavier outer test-mass. Indeed the ratio area over mass is a driving factor of the error budget.

In terms of the Eötvös parameter, these results are statistically compatible with a null value: all absolute values are smaller than 2σ except for sessions 294 and 380 where they exceed 3σ but remain below 5σ. Accounting for systematic errors (see table 10) further downplays the significance of these values. The value of σ depends on the segment and in particular of its length but is typically 10^{-14} or smaller for the longest segments of 120 orbits.

5.5. Robustness tests

We conducted a series of complementary analyses to test the robustness of our results and quantify how they are impacted by the settings of the analysis and by some terms considered as negligible according to the specifications. For reason of CPU time, all these tests have been performed using the ADAM software.

5.5.1. Impact of the frequency bandwidth. As explained above and in [47], we select a frequency band around f_{EP} and $2f_{\mathrm{EP}}$ to compute the parameters of the model using a weighted least-squares estimation in the frequency domain; we have checked that dividing the used bandwidth around these frequencies by a factor 4 does not change the values of the parameters by more than a few percent (a few 10^{-16} for δ_x); however the associated standard deviation can change by up to 10% which is understandable because if we reduce the number of data (in the frequency domain) its estimation is less precise. In the same spirit, if we no longer use a selected band of frequency but the whole domain, the parameters are not noticeably modified and the estimated standard deviation generally increases from 10 to 20%; indeed, in this case, we could have unmodeled high frequency effects which contribute to increasing the standard deviation. This is also probably why σ estimated by M-ECM (which implicitly uses the whole frequency domain) is generally slightly larger than the one obtained with ADAM.

5.5.2. Impact of the degree of the fitted polynomial. In the actual analysis, we estimate a polynomial of degree 3 in order to absorb the low frequency trend due to temperature variations [47]. We have also tested other degrees, from 1 to 5. While there is no significant modification with degrees 2, 4 or 5 the standard deviation is slightly increased when using a degree 1.

We have also compared different strategies to estimate the coefficients of the polynomial:

(a) A prior estimation (without weighting) in the time domain,

(b) A prior estimation (with weighting according to the estimated PSD of the noise) in the Fourier domain (using the whole spectrum),

(c) Estimation at the same time as the other parameters using the whole spectrum but the frequency bands around f_{EP} and $2f_{\mathrm{EP}}$.

The final results for the estimated Eötvös parameter and the associated standard deviation are fully equivalent: the differences are much smaller than the error and than our ideal objective of 10^{-15}.

5.5.3. Impact of the angular velocity and angular acceleration. Angular velocity (included in the matrix [**S**]) and angular acceleration are used in the actual analysis to correct the measured linear acceleration from gradient effects due to offcentring $\overrightarrow{\Delta}$. The stability of the attitude of MICROSCOPE has been specified to limit the effect of this gradient at the $f_{\rm EP}$ and $2f_{\rm EP}$ frequencies. To check that this is indeed the case, we have analysed the data without these corrections. Again, this resulted in negligible changes in the estimated values of the Eötvös parameter, less than 5% of its standard deviation except for 3 segments for which this was about 10%.

5.5.4. Impact of the residual variations of position of the test masses. Fundamentally, the full dynamical equation of the test masses includes terms taking into account the motion of the masses with respect to the satellite (see equation (1) in [38]): the kinematic acceleration $\ddot{\vec{\Delta}}$ and the Coriolis acceleration $2\vec{\Omega} \times \dot{\vec{\Delta}}$. Since the principle of the accelerometer is to nullify the displacement of the test masses, these terms are not included in the actual analysis. Nevertheless position measurements are available in the telemetry data. Due to the limited quantity of data that can be transmitted on ground, they are sampled at 1 Hz. Thus, in order to compute the previous terms and correct the electrostatic acceleration we have both interpolated the position at 4 Hz and derived them numerically. The modification of the estimated δ_x amounts to a few percents of its standard deviation.

5.6. Testing the whole process with a simulated EP signal

It is essential to check that the global analysis process, including the detection-elimination of the glitches, is able to preserve and retrieve a potential EP-violation signal. To this aim we added a fake signal to the real measurements before any preprocessing and applied our complete chain of analysis going from the masking described in section 4.2.2 to the estimation of the Eötvös parameter as explained above. More precisely, we conducted two series of tests with two levels of added simulated signal: one corresponding to an Eötvös parameter of 3.4×10^{-14} (which is large compared to the objective of MICROSCOPE) and a second corresponding to an Eötvös parameter of 3.4×10^{-15} (which is more or less the limit of detection for SUEP as confirmed by equation (19)). We then performed the analysis of these data and subtracted the Eötvös parameter estimated with the original data. Tables 8 and 9 show these differences which can be compared to the simulated signal. The standard deviations corresponding to these analyses are very close to those of the initial analysis (tables 6 and 7) and are not repeated here. Tables 8 and 9 show also the bias (i.e. the difference quoted above minus the simulated value) divided by the standard deviation of the Eötvös parameter. The absolute value of this ratio is smaller than 2% for all segments of the SUEP and for most of the segments of the SUREF. The worst case is for segment 778-1 (SUREF) with a relative error of 18% for a fake signal of 3.4×10^{-14}; but even in this case the absolute error is smaller than 10^{-15}.

5.7. Systematic errors

The evaluation of systematic errors is a major topic addressed during the preparation and specification of the mission and also since the end of the mission in 2018 using the actual measurements and characterisation of the experiment. Systematic errors are estimated in detail in [43] (see table 15 therein for a summary). They can be divided into seven main contributors, the maximum impact of each was estimated at the $f_{\rm EP}$ frequency.

(a) Residual gravitational effects either due to imperfect knowledge of the Earth's gravity field and of the position and orientation of the instrument, or due to local effects coming from

Table 8. Estimation of a simulated fake EP signal to test the robustness of the analysis process for SUREF.

	Fake Eötvös parameter 3.40×10^{-15}		Fake Eötvös parameter 34.00×10^{-15}	
Segment number	Estimated ($\times 10^{15}$)	Bias/σ	Estimated ($\times 10^{15}$)	Bias/σ
120-1	3.41	0.00	34.01	0.00
120-2	3.45	0.01	34.00	0.00
174	3.40	0.00	33.99	0.00
176	3.44	0.01	34.04	0.01
294*	3.34	−0.03	33.97	−0.01
376-1	3.46	0.01	33.93	−0.01
376-2	3.31	−0.02	33.96	−0.01
380-1*	3.41	0.00	34.07	0.03
380-2*	3.23	−0.06	33.81	−0.07
452	3.43	0.01	34.04	0.01
454	3.37	−0.01	33.98	−0.01
778-1	3.39	0.00	33.18	−0.18
778-2	3.44	0.01	34.09	0.02

Table 9. Estimation of a simulated fake EP signal to test the robustness of the analysis process for SUEP.

	Fake Eötvös parameter 3.40×10^{-15}		Fake Eötvös parameter 34.00×10^{-15}	
Segment number	Estimated ($\times 10^{15}$)	Error/σ	Estimated ($\times 10^{15}$)	Error/σ
210*	3.22	−0.01	34.14	0.01
212*	3.36	0.00	33.95	0.00
218*	3.51	0.01	34.13	0.02
234*	3.39	0.00	33.85	−0.02
236*	3.41	0.00	33.99	0.00
238*	3.42	0.00	34.00	0.00
252*	3.37	0.00	33.97	0.00
254*	3.37	0.00	33.98	0.00
256*	3.51	0.01	33.88	−0.02
326-1*	3.42	0.00	34.23	0.02
326-2*	3.33	0.00	33.77	−0.02
358*	3.41	0.00	34.06	0.01
402	3.54	0.00	33.93	0.00
404*	3.38	0.00	34.01	0.00
406*	3.37	0.00	34.00	0.00
438	3.45	0.00	34.19	0.01
442	3.45	0.00	34.02	0.00
748	3.40	0.00	34.08	0.00
750*	3.95	0.01	34.66	0.02

the satellite or the instrument itself (first three rows of table 15 in [43]). Those residuals are due either to errors in the correction of the Earth's gravity gradients after calibration of the offcentrings, or to errors in the estimation of local gravity fluctuations (performed with finite element analysis before the launch).

(b) Clock errors: the most stringent requirement in terms of time stamping comes from the need to compute the gravity gradient tensor with the correct position and orientation of the satellite, in order to correct the effects of the gravity gradient. Since the dominant contributions of the gravity gradient are at DC and $2f_{\rm EP}$ this contrains the absolute clock errors at $f_{\rm EP}$ and $3f_{\rm EP}$[24]; they have been specified to be smaller than 1 ms, leading to an error smaller than 2×10^{-16} ms^{-2} on the gravity gradient effect. The maximum effective errors were 2.1 μs and 0.7 μs respectively at $f_{\rm EP}$ and $3f_{\rm EP}$ in inertial pointing and even smaller in spin mode. The maximum total error (including bias, harmonic errors and drift) with respect to UTC was specified to be smaller than 50 ms and was always smaller than 41.3 ms. In order to limit the effect of the drift, the on-board clock was regularly synchronized (outside the scientific sessions).

(c) Uncorrected inertial effects due to imperfect estimation of angular velocity and angular acceleration. This contribution is computed via the estimated performance of the DFACS [37, 45], which allows estimation of the angular error contribution at $f_{\rm EP}$ for each session.

(d) Time variation of the instrument parameters (satellite pointing, common mode test-mass alignment and angular-to-linear acceleration couplings). These variations are not due to temperature variation at $f_{\rm EP}$ (which are taken into account in the thermal effect below). They are computed by combining the satellite alignment variation issued from the DFACS control with the residual continuous differential acceleration at the differential measurement output.

(e) The drag-free control residuals. In complementarity with the previous item, instrument parameters (estimated from calibration sessions) are assumed constant during a given session. The DFACS-related systematic error is computed from the residual accelerations at $f_{\rm EP}$ issued from the DFACS performance report.

(f) Magnetic effects. They are computed with a finite element model partially adjusted to measurements made on the instrument magnetic shielding and based on the knowledge of all electronics units characteristics.

(g) Thermal effects induced by the tiny variations of temperature at the $f_{\rm EP}$ frequency. These are the major contributors to the systematic error budget. Their estimation results from a detailed analysis of the instrument and of the satellite thermal behaviours. It should be noted that the temperature variations at $f_{\rm EP}$ in the SU were reduced to a fraction of μK in the worst case and helped to strongly reduce the thermal contributor compared to that of [39].

(h) Non-linearities, described by a quadratic term in the measurement equation. Note that equation (5) ignores this term because it is not used in the data process but only in *a posteriori* analysis of errors [43].

According to [43], the last two terms are dozens of times larger than the others for SUEP. In the case of SUREF, non-linear effects are smaller and thermal effects dominate all the others. These various effects are unlikely to be correlated and are quadratically added at the end.

The main systematics (in terms of Eötvös parameter) as well as their total contribution $\mathcal{S}_l$ for each segment l are summarised in tables 10 and 11. Following [43], $\mathcal{S}_l$ is computed according to $\mathcal{S}_l = \frac{1}{g}\sqrt{\sum_k \left(\Gamma^{(d)}_{k,l}\right)^2}$ where $\Gamma^{(d)}_{k,l}$ are the maximum systematic errors of acceleration corresponding to each source k, and $g \simeq 7.9 \times 10^{-15}$ ms^{-2} is the gravity acceleration for MICROSCOPE. The standard deviation issued from the least-squares regression with ADAM

[24] The combinations $f_{\rm EP}$ with DC, $f_{\rm EP}$ with $2f_{\rm EP}$ and $3f_{\rm EP}$ with $2f_{\rm EP}$ can all produce effects at $f_{\rm EP}$.

Table 10. Main systematic effects for SUREF sessions: thermal effects and non-linear (quadratic) effects which are dominant and total computed as the quadratic sum of all effects detailed in [43]. The last column repeats, for comparison, the stochastic error obtained with ADAM presented in table 6. All values are given in equivalent of 10^{-15} for the Eötvös parameter (for example systematic errors in ms^{-2} have been divided by 10^{-15} g $\simeq 7.9 \times 10^{-15}$ ms^{-2}). Segments marked by an asterisk * correspond to sessions with spin V3 and the others to sessions with spin V2.

Segment number	Thermal effect	Quadratic effect	Total systematics	Stochastic error
120-1	5.4	0.1	5.4	15.9
120-2	5.4	0.1	5.4	8.3
174	4.0	0.4	4.0	4.2
176	4.7	0.4	4.7	4.5
294*	2.0	0.4	2.0	2.1
376-1	3.7	0.4	3.7	6.5
376-2	3.7	0.4	3.7	5.8
380-1*	0.7	0.4	0.8	2.4
380-2*	0.7	0.4	0.8	2.8
452	2.1	0.5	2.1	4.1
454	2.7	0.4	2.7	2.8
778-1	3.0	0.3	3.0	4.7
778-2	3.0	0.3	3.0	5.5

Table 11. Same as table 10 but for SUEP.

Segment number	Thermal effect	Quadratic effect	Total systematics	Stochastic error
210*	1.7	0.8	1.8	13.1
212*	0.6	0.9	1.0	11.9
218*	0.8	0.7	1.1	8.1
234*	0.8	0.7	1.0	8.3
236*	1.0	0.7	1.2	6.6
238*	1.0	0.7	1.2	6.4
252*	0.8	0.7	1.1	7.3
254*	1.3	0.8	1.5	7.0
256*	0.8	0.7	1.1	7.4
326-1*	0.8	1.3	1.6	9.6
326-2*	0.8	1.3	1.6	13.5
358*	0.8	0.7	1.1	10.9
402	7.3	0.7	7.3	43.6
404*	0.7	0.7	1.0	6.7
406*	3.1	0.8	3.2	14.9
438	5.4	0.6	5.5	24.6
442	7.2	0.6	7.3	19.1
748	7.2	0.6	7.3	24.6
750*	7.2	0.7	7.3	38.4

presented in section 5.4 is also recalled for comparison. This shows that stochastic errors are clearly dominant for SUEP and marginally dominant for SUREF.

6. Estimation of the Eötvös parameter using the combination of sessions

The results from the analysis of the individual segments show that stochastic errors are larger than systematic errors. In order to improve the signal to noise ratio we have gathered, in a global analysis, all segments included in tables 6 and 7. The polynomial coefficients α_j and the offcentring Δ'_x and Δ'_z are specific to each segment, but the parameters δ_x and δ_z are common. This has been achieved in the Fourier domain (with the ADAM software) as described in [47] and recalled in appendix A. Due to excessively large gaps between segments, we can not simply cumulate the corresponding measurements without being overwhelmed by leakage effects, even when applying dedicated algorithms like M-ECM. Instead, we apply a DFT to each segment and cumulate the transformed equations in the Fourier domain as detailed in appendix A.

As a result we get for SUREF

$$\delta_x = (0.0 \pm 1.1) \times 10^{-15}, \tag{13}$$

and for SUEP

$$\delta_x = (-1.5 \pm 2.3) \times 10^{-15}, \tag{14}$$

where the errors given above are statistical errors at 1 σ. As was to be expected from the results of individual segments, SUEP suffers from a larger statistical error than SUREF despite the larger number of segments used in the combined solution.

In the same spirit, we have analysed the cumulated segments containing fake signals as described in section 5.6. For the SUREF the estimated increment on the Eötvös ratio is 3.38×10^{-15} when 3.40×10^{-15} was simulated and 34.01×10^{-15} when 34.00×10^{-15} was simulated. For the SUEP the results are respectively 3.37×10^{-15} and 33.99×10^{-15}.

Figure 9 shows the histograms of the weighted residual accelerations in the frequency band around f_{EP} after estimation of the parameters and subtraction of the model (7). We have checked that they are compatible with a Gaussian statistics. Note that if instead of gathering all measurements to provide the global solution, we compute the weighted mean of the solutions for individual segments l as

$$\delta_{x,\mathrm{M}} = \frac{\sum_l \frac{\delta_{x,l}}{\sigma_l^2}}{\sum_l \frac{1}{\sigma_l^2}}, \tag{15}$$

and the associated variance

$$\sigma_{\mathrm{M}}^2 = \frac{1}{\sum_l \frac{1}{\sigma_l^2}}, \tag{16}$$

we get the very similar results $\delta_{x,\mathrm{M}} = (-0.4 \pm 1.1) \times 10^{-15}$ for SUREF and $\delta_{x,\mathrm{M}} = (-1.8 \pm 2.2) \times 10^{-15}$ for SUEP. This is expected if the observations of the different segments are sufficiently independent. The same weighting is used to combine systematic errors $\mathcal{S}_l$ associated to individual segments (given in column 4 of tables 10 and 11) in order to get the systematic error $\mathcal{S}_{\mathrm{M}}$ associated to the global solution [43]:

$$\mathcal{S}_{\mathrm{M}} = \frac{\sum_l \frac{\mathcal{S}_l}{\sigma_l^2}}{\sum_l \frac{1}{\sigma_l^2}}, \tag{17}$$

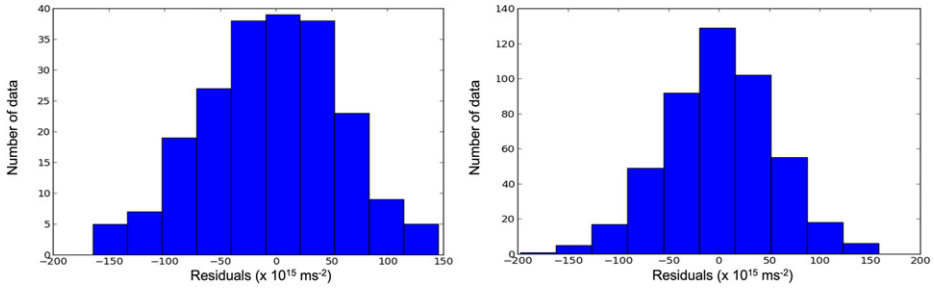

Figure 9. Histograms of the residuals, in the frequency band around $f_{\rm EP}$, of the measured acceleration after fitting of the model (7) in the Fourier domain for SUREF (left panel) and SUEP (right).

which leads to the systematics $\mathcal{S}_{\rm M} = 2.3 \times 10^{-15}$ for SUREF and $\mathcal{S}_{\rm M} = 1.5 \times 10^{-15}$ for SUEP.

Putting all together, and remembering that the conventional Eötvös parameter η can be practically identified to the parameter δ_x measured in this experiment, we end up for SUREF with

$$\eta(\text{Pt,Pt}) \simeq \delta(\text{Pt,Pt}) = [0.0 \pm 1.1(\text{stat}) \pm 2.3(\text{syst})] \times 10^{-15} \text{ at } 1\sigma. \tag{18}$$

As a null Eötvös parameter is expected in this case, this result gives a good indication that there is no important anomaly in the whole chain going from the measurements to the analysis and including the modelling.

We finally get for SUEP

$$\eta(\text{Ti,Pt}) \simeq \delta(\text{Ti,Pt}) = [-1.5 \pm 2.3(\text{stat}) \pm 1.5(\text{syst})] \times 10^{-15} \text{ at } 1\sigma. \tag{19}$$

This final result indicates that there is no visible violation of the WEP in the full data of the MICROSCOPE mission.

7. Conclusion

We have analysed the measurements provided by the payload T-SAGE flown on the MICROSCOPE satellite: these are the differences of accelerations of two test-masses made of the same material (PtRh10) for the sensor SUREF and two test-masses made of different material (PtRh10 for the inner mass, Ti alloys for the outer mass) for the sensor SUEP. This involves 13 segments (i.e. sequences of continuous measurements sampled at 4 Hz) totalling 598 orbits for SUREF and 19 segments totalling 1362 orbits for SUEP. This represents accumulated free falls in the Earth's gravity field of about 41 days for SUREF and 94 days for SUEP. In the data analysis we have compared the measurements to a model including many effects, in particular those of the gravity gradients and of the gradient of inertia due to the tiny difference of positions of the test masses, together with a hypothetical signal of violation of the EP. Before these computations we have corrected and calibrated the instrumental parameters estimated during the dedicated calibration sessions. We have moreover detected and discarded the measurements affected by glitches. As this process breaks the regularity of the sampling, we had to use appropriate algorithms in order to prevent the effects of leakage.

In a first step the analysis has been performed separately on single segments using two different methods: M-ECM operating in the time domain which has been designed for optimal estimation using irregularly sampled data and is also able to reconstruct the most likely data where they are missing, and ADAM operating in the frequency domain. The two methods give

consistent estimates of the parameter $\delta = \eta + \mathcal{O}(\eta^2)$. In particular, the values estimated with the SUEP accelerometer on the different segments are consistent with 0 at less than 2σ for most of them, 2.2σ for one of them. This distribution is compatible with Gaussian statistics. The value of σ depends on the segment and in particular on its length but is typically 10^{-14} or smaller for the longest segments of 120 orbits. The systematic errors, analysed in [43] for the same segments, are significantly smaller.

In a final step we have gathered the data coming from all the segments for each SU in a single analysis, with the aim of obtaining the best signal to noise ratio from our full set of data. This led again to no detection of violation of the WEP since we obtained $\eta(\mathrm{Ti},\mathrm{Pt}) = [-1.5 \pm 2.3(\mathrm{stat}) \pm 1.5(\mathrm{syst})] \times 10^{-15}$ for the SUEP. The result obtained for the SUREF, $\eta(\mathrm{Pt},\mathrm{Pt}) = [0.0 \pm 1.1(\mathrm{stat}) \pm 2.3(\mathrm{syst})] \times 10^{-15}$ confirmed the absence of bias in the whole analysis, a null value being expected in this case. It is common [29, 56] to add quadratically the statistical and systematic errors. Doing this we conclude that the MICROSCOPE experiment does not see evidence for any difference of free fall between titanium and platinum test masses at a level of sensitivity of 2.7×10^{-15}. This represents an improvement of almost two orders of magnitude with respect to the constraint before the launch of MICROSCOPE.

Although this new upper bound on the WEP allows for improved bounds on beyond-GR models (see e.g. [40–42] for bounds obtained after the first MICROSCOPE results [35, 39]), the challenges faced by fundamental physics remain as pressing as ever and call for still more precise experiments. New tests in space could improve MICROSCOPE's state-of-the art measurement by two orders of magnitude in the next decades [57].

Acknowledgments

The authors express their gratitude to all the different services involved in the mission partners and in particular CNES, the French space agency in charge of the satellite. This work is based on observations made with the T-SAGE instrument, installed on the CNES-ESA-ONERA-CNRS-OCA-DLR-ZARM MICROSCOPE mission. ONERA authors' work is financially supported by CNES and ONERA fundings. Authors from OCA, Observatoire de la Côte d'Azur, have been supported by OCA, CNRS, the French National Centre for Scientific Research, and CNES. ZARM authors' work is supported by the DLR, German Space Agency, with funds of the BMWi (FKZ 50 OY 1305 and FKZ 50 LZ 1802) and by the Deutsche Forschungsgemeinschaft DFG (LA 905/12-1). The authors would like to thank the Physikalisch-Technische Bundesanstalt Institute in Braunschweig, Germany, for their contribution to the development of the test-masses with funds of CNES and DLR.

Data availability statement

The data that support the findings of this study will be openly available following an embargo at the following URL:https://cmsm-ds.onera.fr/. Data will be available from 01 October 2022.

Appendix A. Least-squares regression in the Fourier domain

We recall here the main steps of the procedure described in [47].

The set of measurement equation (7) applied at each time of measurement t_i can be written as a linear system

$$\mathbf{Y} = [\mathbf{A}]\boldsymbol{\theta} + \mathbf{n}, \tag{A.1}$$

where $\mathbf{Y}$ is the vector of N measurements, $\boldsymbol{\theta}$ is a vector of q unknown parameters to estimate (i.e. α_j, δ_x, δ_z, Δ'_x and Δ'_z), $[\mathbf{A}]$ is the design matrix and $\mathbf{n}$ is the noise vector. Since the model is linear with respect to the estimated parameters, the columns of $[\mathbf{A}]$ simply correspond to the signal associated to each parameter, sampled at the epochs of the measured acceleration.

The N measurements are assumed to be regularly sampled at a frequency $f_{\rm e}$ over a duration T. In order to solve the problem in the Fourier domain, we take the Fourier transform of equation (7). To this aim, we make use of the DFT operator $[\mathbf{F}]$. The DFT operator being unitary, the signal energy content is preserved by the transformation. The new system can be simply written

$$\hat{\mathbf{Y}} = \left[\hat{\mathbf{A}}\right]\boldsymbol{\theta} + \hat{\mathbf{n}}. \tag{A.2}$$

The original quantities being real, the new system can be reduced to N useful real equations. These new equations can be grouped by pair (related to real and imaginary parts of the DFT), corresponding to frequencies $f_k = \frac{k}{T},\ k = 1 \cdots \lfloor \frac{N-1}{2} \rfloor$.

Since each measurement projected in the Fourier domain can be associated to a discrete frequency, the corresponding weight is

$$w(f_k) = \frac{1}{\sqrt{\gamma(f_k)}}, \tag{A.3}$$

where $\gamma(f_k)$ is the PSD of the noise at the frequency f_k.

As shown in reference [38], the MICROSCOPE mission was designed to concentrate useful signal at specific frequencies (i.e., gravity acceleration peaks at $f_{\rm EP}$, the gravity gradient signal at $2f_{\rm EP}$ and calibration signals at $f_{\rm cal}$). This is so true in the real data that a very simple analysis such as synchronous detection could lead to reasonable results. However, we use a more flexible method: we limit our least-squares inversion to the bands of frequency containing the relevant signals. In practice, this is equivalent to extracting a subsystem of equation (A.2) by selecting the relevant equations to get the truncated system

$$\left[\hat{\mathbf{A}}_r\right]\boldsymbol{\theta} + \hat{\mathbf{n}} = \hat{\mathbf{Y}}_r. \tag{A.4}$$

One can cumulate the data from disjoint segments by just gathering the corresponding matrices and vectors:

$$\left[\hat{\mathbf{A}}\right] = \begin{bmatrix} \left[\hat{\mathbf{A}}_1\right] \\ \left[\hat{\mathbf{A}}_2\right] \\ \vdots \\ \left[\hat{\mathbf{A}}_m\right] \end{bmatrix}, \qquad \hat{\mathbf{n}} = \begin{bmatrix} \hat{\mathbf{n}}_1 \\ \hat{\mathbf{n}}_2 \\ \vdots \\ \hat{\mathbf{n}}_m \end{bmatrix}, \qquad \hat{\mathbf{Y}} = \begin{bmatrix} \hat{\mathbf{Y}}_1 \\ \hat{\mathbf{Y}}_2 \\ \vdots \\ \hat{\mathbf{Y}}_m \end{bmatrix} \tag{A.5}$$

where m is the number of sessions considered. Then an appropriately weighted least-squares technique can be used to solve for the concatenated system, the weight associated to each frequency for each segment being chosen according to (A.3).

In the simplest cases, all parameters are common to all segments. If some parameters are specific to each segment (as could be the case for the polynomial coefficients or for the offcentring), the corresponding column of the design matrix related to the other segments is simply set to zero.

List of acronyms and abbreviations

ADAM	Accelerometric data analysis for MICROSCOPE
DC	Direct continuous
DFACS	Drag-free and attitude control system
DFT	Discrete Fourier transform
EP	Equivalence principle
IS	Inertial sensor
M-ECM	Modified expectation conditional maximisation
MLI	Multi-layer insulation
PSD	Power spectrum density
SU	Sensor unit
SUEP	Sensor unit for the equivalence principle test
SUREF	Sensor unit for reference
T-SAGE	Twin space accelerometer for gravity experiment is the name of the payload
UTC	Universal time coordinated
WEP	Weak equivalence principle

ORCID iDs

Gilles Métris https://orcid.org/0000-0002-3851-3237
Manuel Rodrigues https://orcid.org/0000-0002-8573-8068
Joel Bergé https://orcid.org/0000-0002-7493-7504
Stefanie Bremer https://orcid.org/0000-0002-8091-0121
Meike List https://orcid.org/0000-0002-5268-5633
Sandrine Pires https://orcid.org/0000-0002-0249-2104
Timothy Sumner https://orcid.org/0000-0002-3572-600X

References

[1] Einstein A 1908 Über das Relativitätsprinzip und die aus demselben zogenen Folgerungen *Jahrbuch der Radioaktivität und Elektronik* vol 4 (Leipzig: S. Hirzel)
[2] Einstein A 1916 *Ann. Phys.* **354** 769–822
[3] Will C M 2014 *Living Rev. Relativ.* **17** 4
[4] Zyla P A *et al* (Particle Data Group) 2020 *Prog. Theor. Exp. Phys.* **2020** 083C01
[5] Zwicky F 1933 Die Rotverschiebung von extragalaktischen Nebeln *Helv. Phys. Acta* **6** 110–27 This DOI corresponds to an english transaltion of 2009
[6] Rubin V C and Ford W K J Jr 1970 *Astrophys. J.* **159** 379
[7] Riess A G *et al* 1998 *Astron. J.* **116** 1009–38
[8] Perlmutter S *et al* 1999 *Astrophys. J.* **517** 565–86
[9] Fayet P 1977 *Phys. Lett.* B **69** 489–94
[10] Kaluza T 1921 Zum Unitätsproblem in der Physik. *Sitzungsberichte Preuss. Akad. Wiss.* **1921** 966–72 This DOI corresponds to an englisah translation of 2018
[11] Klein O 1926 *Z. Phys.* **37** 895–906
[12] Scherk J and Schwarz J H 1974 *Nucl. Phys.* B **81** 118–44
[13] Vainshtein A I 1972 *Phys. Lett.* B **39** 393–4
[14] Damour T and Polyakov A M 1994 *Nucl. Phys.* B **423** 532–58
[15] Damour T and Nordtvedt K 1993 *Phys. Rev. Lett.* **70** 2217–9
[16] Khoury J and Weltman A 2004 *Phys. Rev.* D **69** 044026

[17] Khoury J and Weltman A 2004 *Phys. Rev. Lett.* **93** 171104
[18] Babichev E, Deffayet C and Ziour R 2009 *Int. J. Mod. Phys.* D **18** 2147–54
[19] Hinterbichler K and Khoury J 2010 *Phys. Rev. Lett.* **104** 231301
[20] Brax P, Burrage C and Davis A-C 2013 *J. Cosmol. Astropart. Phys.* JCAP01(2013)020
[21] Burrage C and Sakstein J 2018 *Living Rev. Relativ.* **21** 1
[22] Damour T, Piazza F and Veneziano G 2002 *Phys. Rev. Lett.* **89** 081601
[23] Damour T and Donoghue J F 2010 *Phys. Rev.* D **82** 084033
[24] Fayet P 1990 *Nucl. Phys.* B **347** 743–68
[25] Fayet P 2017 *Eur. Phys. J.* C **77** 53
[26] Eötvös RV, Pekár D and Fekete E 1922 Beiträge zum Gesetze der Proportionalität von Trägheit und Gravität *Ann. Phys.* **373** 11–66
[27] Fischbach E, Sudarsky D, Szafer A, Talmadge C and Aronson S H 1986 *Phys. Rev. Lett.* **56** 3–6
[28] Schlamminger S, Choi K Y, Wagner T A, Gundlach J H and Adelberger E G 2008 *Phys. Rev. Lett.* **100** 041101
[29] Wagner T A, Schlamminger S, Gundlach J H and Adelberger E G 2012 *Class. Quantum Grav.* **29** 184002
[30] Williams J G, Turyshev S G and Boggs D H 2012 *Class. Quantum Grav.* **29** 184004
[31] Viswanathan V, Fienga A, Minazzoli O, Bernus L, Laskar J and Gastineau M 2018 *Mon. Not. R. Astron. Soc.* **476** 1877–88
[32] Chapman P K and Hanson A J 1970 *Proc. Conf. on Experimental Tests of Gravitation Theories JPL TM 33-499* ed R W Davies p 228
[33] Everitt C W F, Damour T, Nordtvedt K and Reinhard R 2003 *Adv. Space Res.* **32** 1297–300
[34] Liorzou F *et al* 2022 *Class. Quantum Grav.* **39** 204002
[35] Touboul P *et al* 2019 *Class. Quantum Grav.* **36** 225006
[36] Touboul P, Rodrigues M, Métris G and Tatry B 2001 *C. R. Acad. Sci., Paris IV* **2** 1271
[37] Rodrigues M *et al* 2022 *Class. Quantum Grav.* **39** 204003
[38] Touboul P *et al* 2022 *Class. Quantum Grav.* **39** 204001
[39] Touboul P *et al* 2017 *Phys. Rev. Lett.* **119** 231101
[40] Bergé J, Brax P, Métris G, Pernot-Borràs M, Touboul P and Uzan J-P 2018 *Phys. Rev. Lett.* **120** 141101
[41] Fayet P 2018 *Phys. Rev.* D **97** 055039
[42] Fayet P 2019 *Phys. Rev.* D **99** 055043
[43] Rodrigues M *et al* 2022 *Class. Quantum Grav.* **39** 204006
[44] Touboul P, Métris G, Rodrigues M, Bergé J, Chhun R, Foulon B and Robert A 2022 *Class. Quantum Grav.* **39** 200401
[45] Rodrigues M *et al* 2022 *Class. Quantum Grav.* **39** 204004
[46] Chhun R *et al* 2022 *Class. Quantum Grav.* **39** 204005
[47] Bergé J *et al* 2022 *Class. Quantum Grav.* **39** 204007
[48] Bergé J, Baghi Q, Robert A, Rodrigues M, Foulon B, Hardy E, Métris G, Pires S and Touboul P 2022 *Class. Quantum Grav.* **39** 204008
[49] Flury J, Bettadpur S and Tapley B D 2008 *Adv. Space Res.* **42** 1414–23
[50] Akhlaghi M and Ichikawa T 2015 *Astrophys. J. Suppl.* **220** 1
[51] Baghi Q, Métris G, Bergé J, Christophe B, Touboul P and Rodrigues M 2015 *Phys. Rev.* D **91** 062003
[52] Baghi Q, Métris G, Bergé J, Christophe B, Touboul P and Rodrigues M 2016 *Phys. Rev.* D **93** 122007
[53] Dempster A P, Laird N M and Rubin D B 1977 *J. R. Stat. Soc.* B **39** 1–22
[54] Meng X-L and Rubin D B 1993 *Biometrika* **80** 267–78
[55] Zhu Z and Wakin M B 2017 On the asymptotic equivalence of circulant and toeplitz matrices *IEEE Trans. Inf. Theory* **63** 2975–92
[56] Joint Committee for Guides in Metrology 2008 *Evaluation of Measurement Data; the Role of Measurement Uncertainty in Conformity Assessment* BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML
[57] Battelier B *et al* 2021 *Exp. Astron.* **51** 1695–736