Editors' Suggestion | Featured in Physics

Beyond the Required LISA Free-Fall Performance: New LISA Pathfinder Results down to 20 μHz

M. Armano,[1] H. Audley,[2] J. Baird,[3] P. Binetruy,[4,*] M. Born,[2] D. Bortoluzzi,[5] E. Castelli,[6] A. Cavalleri,[7] A. Cesarini,[8] A. M. Cruise,[9] K. Danzmann,[2] M. de Deus Silva,[1] I. Diepholz,[2] G. Dixon,[9] R. Dolesi,[6] L. Ferraioli,[10] V. Ferroni,[6] E. D. Fitzsimons,[11] M. Freschi,[1] L. Gesa,[12] F. Gibert,[6] D. Giardini,[10] R. Giusteri,[6] C. Grimani,[8] J. Grzymisch,[13] I. Harrison,[14] G. Heinzel,[2] M. Hewitson,[2] D. Hollington,[3] D. Hoyland,[9] M. Hueller,[6] H. Inchauspé,[4] O. Jennrich,[13] P. Jetzer,[15] N. Karnesis,[2] B. Kaune,[2] N. Korsakova,[16] C. J. Killow,[16] J. A. Lobo,[12,*] I. Lloro,[12] L. Liu,[6] J. P. López-Zaragoza,[12] R. Maarschalkerweerd,[14] D. Mance,[10] N. Meshksar,[10] V. Martín,[12] L. Martin-Polo,[1] J. Martino,[4] F. Martin-Porqueras,[1] I. Mateos,[12] P. W. McNamara,[13] J. Mendes,[14] L. Mendes,[1] M. Nofrarias,[12] S. Paczkowski,[2] M. Perreur-Lloyd,[16] A. Petiteau,[4] P. Pivato,[6] E. Plagnol,[4] J. Ramos-Castro,[17] J. Reiche,[2] D. I. Robertson,[16] F. Rivas,[12] G. Russano,[6] J. Slutsky,[18] C. F. Sopuerta,[12] T. Sumner,[3] D. Texier,[1] J. I. Thorpe,[18] D. Vetrugno,[6] S. Vitale,[6] G. Wanner,[2] H. Ward,[16] P. J. Wass,[3,19] W. J. Weber,[6] L. Wissel,[2] A. Wittchen,[2] and P. Zweifel[10]

[1]*European Space Astronomy Centre, European Space Agency, Villanueva de la Cañada, 28692 Madrid, Spain*
[2]*Albert-Einstein-Institut, Max-Planck-Institut für Gravitationsphysik und Leibniz Universität Hannover, Callinstraße 38, 30167 Hannover, Germany*
[3]*High Energy Physics Group, Physics Department, Imperial College London, Blackett Laboratory, Prince Consort Road, London, SW7 2BW, United Kingdom*
[4]*APC, Université Paris Diderot, CNRS/IN2P3, CEA/lrfu, Observatoire de Paris, Sorbonne Paris Cité, France*
[5]*Department of Industrial Engineering, University of Trento, via Sommarive 9, 38123 Trento, and Trento Institute for Fundamental Physics and Application/INFN, Italy*
[6]*Dipartimento di Fisica, Università di Trento and Trento Institute for Fundamental Physics and Application/INFN, 38123 Povo, Trento, Italy*
[7]*Istituto di Fotonica e Nanotecnologie, CNR-Fondazione Bruno Kessler, I-38123 Povo, Trento, Italy*
[8]*DISPEA, Università di Urbino "Carlo Bo," Via S. Chiara, 27 61029 Urbino/INFN, Italy*
[9]*The School of Physics and Astronomy, University of Birmingham, Birmingham, United Kingdom*
[10]*Institut für Geophysik, ETH Zürich, Sonneggstrasse 5, CH-8092, Zürich, Switzerland*
[11]*The UK Astronomy Technology Centre, Royal Observatory, Edinburgh, Blackford Hill, Edinburgh, EH9 3HJ, United Kingdom*
[12]*Institut de Ciències de l'Espai (ICE, CSIC), Campus UAB, Carrer de Can Magrans s/n, 08193 Cerdanyola del Vallès, and Institut d'Estudis Espacials de Catalunya (IEEC), Gran Capità, 2-4, 08034 Barcelona, Spain*
[13]*European Space Technology Centre, European Space Agency, Keplerlaan 1, 2200 AG Noordwijk, Netherlands*
[14]*European Space Operations Centre, European Space Agency, 64293 Darmstadt, Germany*
[15]*Physik Institut, Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland*
[16]*SUPA, Institute for Gravitational Research, School of Physics and Astronomy, University of Glasgow, Glasgow, G12 8QQ, United Kingdom*
[17]*Department d'Enginyeria Electrònica, Universitat Politècnica de Catalunya, and Institut d'Estudis Espacials de Catalunya (IEEC), Gran Capità, 2-4, 08034 Barcelona, Spain*
[18]*Gravitational Astrophysics Lab, NASA Goddard Space Flight Center, 8800 Greenbelt Road, Greenbelt, Maryland 20771, USA*
[19]*Department of Mechanical and Aerospace Engineering, MAE-A, P.O. Box 116250, University of Florida, Gainesville, Florida 32611, USA*

(Received 11 December 2017; revised manuscript received 12 January 2018; published 5 February 2018)

In the months since the publication of the first results, the noise performance of LISA Pathfinder has improved because of reduced Brownian noise due to the continued decrease in pressure around the test masses, from a better correction of noninertial effects, and from a better calibration of the electrostatic force actuation. In addition, the availability of numerous long noise measurement runs, during which no perturbation is purposely applied to the test masses, has allowed the measurement of noise with good statistics down to 20 μHz. The Letter presents the measured differential acceleration noise figure, which is

at (1.74 ± 0.01) fm s$^{-2}/\sqrt{\text{Hz}}$ above 2 mHz and $(6 \pm 1) \times 10$ fm s$^{-2}/\sqrt{\text{Hz}}$ at 20 μHz, and discusses the physical sources for the measured noise. This performance provides an experimental benchmark demonstrating the ability to realize the low-frequency science potential of the LISA mission, recently selected by the European Space Agency.

DOI: 10.1103/PhysRevLett.120.061101

Introduction.—LISA Pathfinder (LPF) [1] is a European Space Agency (ESA) mission dedicated to the experimental demonstration of the free fall of test masses (TMs) as required by LISA [2], the space-based gravitational-wave (GW) observatory just approved by ESA. Such TMs are the reference bodies at the ends of each LISA interferometer arm and need to be free from spurious acceleration, g, relative to their local inertial frame; any stray acceleration competes directly with the tidal deformations caused by GWs. LPF has two LISA TMs at the ends of a short interferometer arm, insensitive to GWs because of the reduced length but sensitive to the differential acceleration, Δg, of the TMs arising from parasitic forces.

LPF was launched on December 3, 2015 and was in science operation from March 1, 2016. Operations ended on June 30, 2017, and the satellite was finally passivated on July 18, 2017. On June 7, 2016, we published [3] the first results on the free fall performance of the LPF test masses. These results showed that the amplitude spectral density (ASD) of Δg was found to be (see Fig. 1 of Ref. [3]) limited by Brownian noise at $S_{\Delta g}^{1/2} = (5.2 \pm 0.1)$ fm s$^{-2}/\sqrt{\text{Hz}}$, for frequencies $1\ \text{mHz} \lesssim f \lesssim 30$ mHz; rising above the Brownian noise floor for frequencies $f \lesssim 1$ mHz, increasing to $\lesssim 12$ fm s$^{-2}/\sqrt{\text{Hz}}$ at $f = 0.1$ mHz; and limited, for $f \gtrsim 30$ mHz, by the interferometer readout noise of $S_x^{1/2} = (34.8 \pm 0.3)$ fm$/\sqrt{\text{Hz}}$, which translates into an effective Δg ASD of $S_x^{1/2}(2\pi f)^2$.

The previously published data referred to the longest uninterrupted stretch of data, of about one week duration, we had measured up to the time of publication. Since that time, several improvements have allowed a significantly better performance, presented in Fig. 1. First, the residual gas pressure has decreased by roughly a factor of 10 since the beginning of operations, as the gravitational reference sensor (GRS) surrounding the TM has been continuously vented to space [3] with a slowly decreasing outgassing rate. Second, a more accurate calculation of the electrostatic actuation force has eliminated a systematic source of low-frequency force noise. Third, another inertial force from the LPF spacecraft rotation has been identified and corrected in the Δg time series. This last effect will be highly suppressed in LISA by the improved rotational spacecraft control. Finally, we have removed, by empirical fitting, a number of well-identified, sporadic (less than one per day) quasi-impulse force events or "glitches" from the data, allowing uninterrupted data series of up to ~18 days duration. This

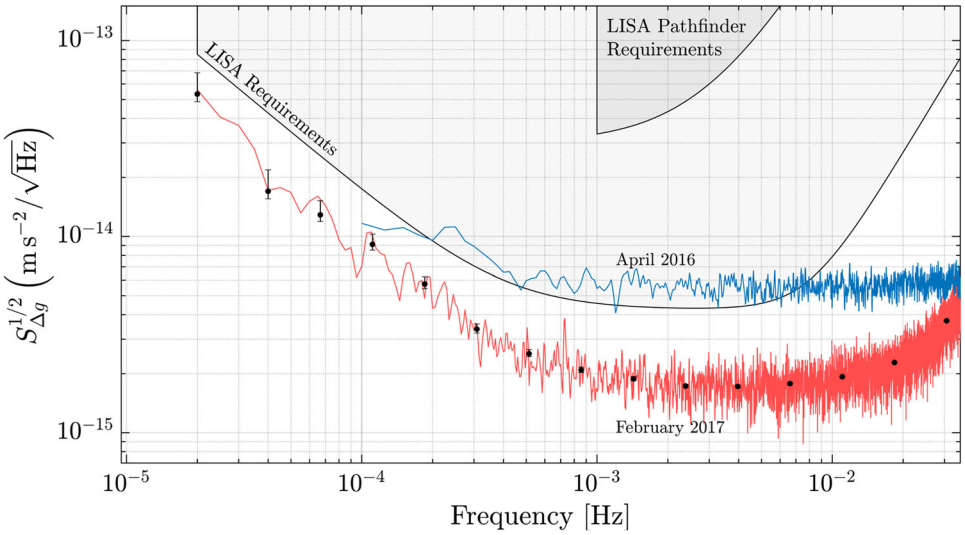

FIG. 1. ASD of parasitic differential acceleration of LPF test masses as a function of the frequency. Data refer to an ~13 day long run taken at a temperature of 11 °C. The red, noisy line is the ASD estimated with the standard periodogram technique averaging over 10, 50% overlapping periodograms each 2×10^5 s long. The data points with error bars are uncorrelated, averaged estimates calculated as explained in the text. For comparison, the blue noisy line is the ASD published in Ref. [3]. Data are compared with LPF requirements [1] and with LISA requirements taken from Ref. [2]. Fulfilling requirements implies that the noise must be below the corresponding shaded area at all frequencies. LISA requirements below 0.1 mHz must be considered just as goals [2].

has allowed us to estimate noise, with good statistics, at significantly lower frequencies than in Ref. [3].

Our Letter addresses these effects individually, to justify the estimate of the acceleration noise ASD relevant to the LISA TM that we show in Fig. 1, with brief comments on the observational possibilities for the LISA observatory, as formulated in Ref. [2]. We also discuss the possible sources of the residual noise, in particular, at a low frequency and in the glitch events.

The LISA Pathfinder instrument and experiment.—The LPF instrument and experimental approach have been described in Refs. [3,4]. In summary, the core instrument, the LISA Technology Package (LTP), consists of two gold-platinum, quasicubic test masses, of size (46.000 ± 0.005) mm and mass $M = (1.928 \pm 0.001)$ kg. These masses are in free fall inside a single spacecraft with their centers separated by a nominal distance of $L = (376.00 \pm 0.05)$ mm along a line that we take as the x axis (see Fig. 1 of Ref. [3]).

The full, multiple degree-of-freedom sensing, dynamics and control of LISA Pathfinder are described in Refs. [3,5], and we recall here only the essential elements of the main x-axis science measurement. The science signal Δg is mainly in the differential displacement of the two TMs along the x axis, $\Delta x \equiv x_2 - x_1$, measured by a dedicated heterodyne laser interferometer [6–8], and in the applied actuation force per unit mass, $g_c(t)$, used to control TM2 to follow TM1, using a slow controller with unity gain around 1 mHz. An additional interferometer readout gives the absolute displacement of the spacecraft with respect to TM1, x_1, which is used for the spacecraft "drag-free" control.

The calculation of the differential force per unit mass, Δg, is given by

$$\Delta g(t) \equiv \Delta\ddot{x}(t) + \omega_2^2 \Delta x(t) + \Delta\omega_{12}^2 x_1(t) - g_c(t) - g_\Omega(t). \tag{1}$$

Force gradients per unit mass, or stiffnesses, $-\omega_1^2$ and $-\omega_2^2$ are defined for the two TMs around their nominal centered positions in terms of their corresponding natural harmonic oscillator angular frequencies. $\Delta\omega_{12}^2 \equiv \omega_2^2 - \omega_1^2$ is the differential stiffness that couples spacecraft motion into Δg. The differential acceleration $\Delta\ddot{x}(t)$ is calculated as the numerical second time derivative of $\Delta x(t)$, while $g_c(t)$ is the time series of the control force applied to TM2. Stiffness terms in Eq. (1) are calculated from the measured displacement time series, $\Delta x(t)$ and $x_1(t)$. g_Ω is the centrifugal force per unit mass, calculated by integrating the average control torques applied to align the two TMs to the spacecraft orientation and from the quasistatic part of the angular velocity derived from the star trackers as explained in Ref. [3].

Acceleration noise data.—The ASD of Δg is estimated during noise-only runs in which no other forces, except for g_c, are applied to the TMs. These noise measurements were interleaved with dedicated measurements of known noise sources and tests of various other hardware and control techniques throughout the mission. Most of g_c is used to compensate the quasistatic differential force that would make the TMs drift apart. The applied values for g_c have been limited by the accuracy of gravitational balancing achieved in LPF [3,9], which has allowed the use of an actuation authority of 50 pN for the TM2 x force, corresponding to applied voltages of the order of 1 V. The required TM ϕ torques, applied with the same X electrodes and thus a source of force noise along the x axis, are also small, allowing authorities of roughly 1 pNm and similar voltage levels. The resulting actuation force noise is thus smaller than foreseen in preflight estimates [5] and roughly 4.5 $\mathrm{fm\,s^{-2}/\sqrt{Hz}}$ at 0.1 mHz, based on dedicated actuation noise experiments performed during the mission.

Following the publication of Ref. [3], the scrutiny of the actuation electronics algorithms revealed a systematic error in the conversion of the commanded peak actuation voltage into the effective rms voltage of the digitized audio waveform. The effect depends in a complicated but systematic—and calculable—fashion on the exact commanded voltages and has been confirmed on an engineering model of the electronics. The resulting error in the calculated force converts a slow drift in the commanded g_c into an additional pseudorandom noise. This exceeds 10 $\mathrm{fm\,s^{-2}/\sqrt{Hz}}$ at 0.1 mHz in the cases of the largest drift and, though a smaller effect in the published results of Ref. [3], still gave a measurable overestimate of the low-frequency noise. The correction of this effect is essential for the new measurements shown here.

LPF measurements are performed in the noninertial reference frame of the spacecraft, whose rotation is controlled using a set of autonomous star trackers with a much larger noise than will be provided by the differential wavefront sensing of distant spacecraft to be used on LISA.

This large rotational noise gives rise to a noisy centrifugal acceleration [3] and to a noisy Euler force connected to the noisy angular acceleration $\dot{\vec{\Omega}}$ of the spacecraft. For any misalignment between the differential interferometer sensitive axis $\hat{n}$ and the vector $\Delta\vec{r} \equiv \vec{r}_2 - \vec{r}_1$ joining the TM centers of mass, the Euler force acquires a projection into the sensitive axis, $\Delta g_E = (\Delta\vec{r} \times \dot{\vec{\Omega}}) \cdot \hat{n}$.

The relevant offset components of $\Delta\vec{r}$ are found by fitting the Δg time series to the relevant spacecraft rotational angular accelerations $\dot{\Omega}_y$ and $\dot{\Omega}_z$, as determined from the applied TM torques. Here we are using the same definition for linear and angular coordinates as in Ref. [3]. The offsets Δy and Δz—expressed in what follows in terms of angular offsets between the inter-TM and interferometer axes, $\delta_y \equiv (\Delta z/L)$ and $\delta_z \equiv (\Delta y/L)$—are well resolved throughout the mission, ranging up to 200 μm displacements and thus angular offsets of up to 0.5 mrad. The

changes in $\Delta\vec{r}$ measured in this way agree with programed offsets in the TM control points. Additionally, the effect is calibrated by the effective rejection of occasional "events" of large spacecraft roll in the Δg data. We note that this geometric Euler effect is indistinguishable, given how it is calculated, with possible torque-to-force actuation cross talk, and we effectively correct for the sum of the two effects.

Our LPF observable, for the differential force per unit mass relevant to the LISA TM acceleration, thus becomes

$$\Delta g(t) \equiv \Delta\ddot{x}(t) + \omega_2^2 \Delta x(t) + \Delta\omega_{12}^2 x_1(t) - g_c(t) - g_\Omega(t) - \delta_y L\dot{\Omega}_y - \delta_z L\dot{\Omega}_z. \quad (2)$$

Finally, we observe in the Δg time series, once properly low-passed to suppress the interferometer noise, clearly identifiable events that we will refer to as glitches in the rest of the Letter. These glitches are observationally indistinguishable from a quasi-impulsive force acting on one of the two TMs, transferring a differential momentum over time spans ranging from seconds to, in rarer cases, hours. Observed glitch amplitudes are as large as $\mathrm{pm/s^2}$, with a typical impulse Δv of the order of 10 pm/s. Over the 76 days of total noise runs we consider in this Letter, we have found a total of 48 glitches. These glitches appear to occur with Poissonian statistics, with an average rate of 0.78 ± 0.02/day. These glitches can be removed from the data as explained in Ref. [10]. Removing the glitches allowed us to generate long uninterrupted Δg data series. In particular, we have four independent noise runs of a duration longer than 9 days and a total of eight with a duration exceeding 5 days (a summary table for these runs can be found in Ref. [11]).

The noise measurements we present in this Letter refer to time series that have been corrected for all the above effects.

The ASD estimate for the data in Fig. 1 has been obtained with two different methods. With the first, we estimate the ASD with the standard periodogram method as described in Ref. [3] and extending now the analysis down to 20 μHz. The second method uses Bayesian inference to estimate the $1-\sigma$, that is, the 68.3% confidence interval for the value of the ASD at selected frequencies. These frequencies are chosen so that the estimates at adjoining frequencies are not significantly correlated. At each of the selected frequencies, the length of the periodograms is adjusted to achieve the maximum averaging and resolution. Details of the method are described in Ref. [12].

Discussion of results.—The ASD of Δg, measured from a single contiguous 13 day period of noise data, measured in February 2017, is presented in Fig. 1 as a benchmark of the LPF differential acceleration noise performance, in comparison with the requirements for LPF and for the proposed LISA mission, as well as the previously published data from Ref. [3].

At frequencies above 10 mHz, the noise is dominated by displacement measurement noise, at a level of roughly 100 $\mathrm{fm}/\sqrt{\mathrm{Hz}}$, an increase over the 35 $\mathrm{fm}/\sqrt{\mathrm{Hz}}$ reported in Ref. [3]. This increase is not attributed to the intrinsic noise in the interferometer readout itself, which has been found to be the same within errors, but rather can be explained by a change in the TM position and attitude control points. This increases the noise associated with correcting the data for cross talk due to satellite motion, using noisier capacitive sensing measurements (this geometric cross talk effect, the resulting raw noise "bulge" above 10 mHz, and its correction are discussed in Ref. [3]). This noise remains, however, negligible compared to the LPF and LISA interferometry requirements.

The dominant noise source in the 1–10 mHz band is Brownian motion of test masses caused by the residual gas surrounding them, as discussed in Ref. [3]. The pressure has since continued to decrease, proportionally to the decreasing outgassing rates in the continuously vented-to-space vacuum chamber housing the TM. In addition, to further reduce the pressure before collecting the data in Fig. 1, the temperature of the entire LTP was reduced, in January 2017, from the 22 °C–23 °C range to 11 °C–12 °C. The Brownian differential acceleration noise $\bar{S}_B$, as estimated by the average power spectral density (PSD) in the minimum noise band from 2 to 4 mHz, decreased by a factor of 0.38 ± 0.03 from the January run to the colder February measurement, implying a reduction in pressure by the same factor if Brownian noise is indeed dominant. Such a change of pressure with the temperature corresponds to an activation energy of the desorbed species $E = (65 \pm 7)$ MJ/kmol, not an unlikely value for water on metals. The overall reduction of pressure from the first long run in March 2016 to the February 2017 run is roughly a factor of 10. If the dominant molecular species is water, then the final average PSD of $\bar{S}_B = [(1.74 \pm 0.05)\ \mathrm{fm\,s^{-2}}/\sqrt{\mathrm{Hz}}]^2$ corresponds to a pressure $P \simeq 1$ μPa [13].

The noise rises above the Brownian noise level at lower frequencies, with a nearly $1/f^2$ dependence in power that becomes dominant below roughly 1 mHz. In parallel with the continuously decreasing Brownian noise, the low-frequency excess noise has remained remarkably stable over a large portion of the mission, including the long February run shown in Fig. 1, allowing averaging with high statistical precision. Figure 2 shows an average of the low-frequency noise in excess of the measured, time-variable "white" Brownian noise floor, for four runs covering 46 days over 7 months from July 2016 to February 2017. The technique for estimating this low-frequency noise excess, and the related uncertainties, is presented in Ref. [14]. A test for the consistency of these data, also discussed in Ref. [12], with the hypothesis that all runs from this group obey the statistics of a single underlying noise spectrum gives a probability larger than 15% at all frequencies studied. A noise projection based on noise sources with well-understood and experimentally anchored models explains a significant, but not dominant, fraction of

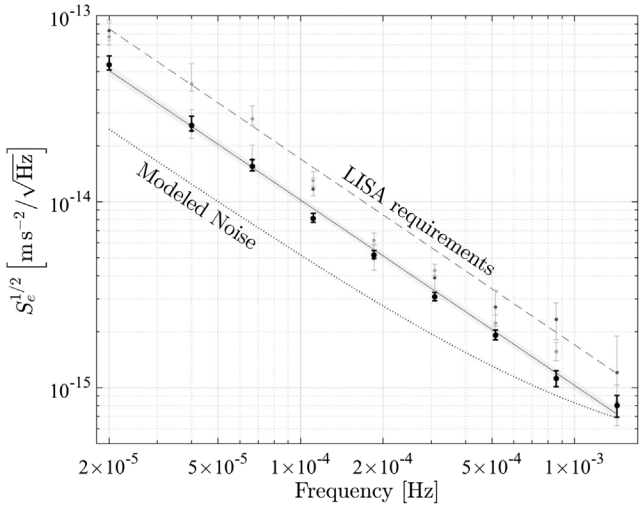

FIG. 2. ASD of excess noise above the Brownian limit. Data and error bars are obtained as explained in the text and in Ref. [12]. Black data refer to all runs between July 2016 and February 2017 (group B). The solid line and accompanying shadowed area represent the best fit of the black data to a power law with its 68% confidence area, following $S_e^{1/2} = [(1.0 \pm 0.3)\ \text{am s}^{-2}\sqrt{\text{Hz}}]f[\text{Hz}]^{(-1.00\pm0.04)}$. The upper dashed straight line is the LISA requirement minus the flat $\sqrt{2} \times 3\ \text{fm s}^{-2}/\sqrt{\text{Hz}}$ requirement dominant at millihertz frequencies. The dotted lower line is our best estimate of the noise produced by modeled contributions. For comparison, the darker gray data in the background refer to the short group of runs taken at the beginning of operations (March–April 2016, group A) and the lighter gray ones to the final short group (May–June 2017, group C) taken immediately after the 0 °C cooldown maneuver and just before the end of the operations.

this low-frequency excess noise. The projected ASD is plotted as the dotted line in Fig. 2. This estimate includes (i) fluctuations in the x/ϕ (force)/(torque) actuation control voltages, as discussed in the previous section (roughly 4.5 fm s$^{-2}/\sqrt{\text{Hz}}$ at 0.1 mHz), (ii) low-frequency electrostatic potential fluctuations interacting with the nonzero TM charge [15], (iii) pressure fluctuations from laser intensity noise (roughly 2 fm s$^{-2}/\sqrt{\text{Hz}}$), and (iv) random charge fluctuations [15], interacting with the residual average stray dc electrostatic fields, which have been measured and compensated to the 5 mV level (well below 1 fm s$^{-2}/\sqrt{\text{Hz}}$). Other known effects, including low-frequency magnetic fluctuations and coupling to the fluctuations of the GRS temperature and temperature gradients, are still under study and are not included in this noise budget. However, based on correlation studies with environmental diagnostics as well as dedicated measurements, these are not expected to contribute significantly to the noise projection, with the possible exception of the mean GRS temperature at the very lowest frequencies studied.

Thus, at the lowest frequencies we observe a definite excess over that explained by our noise model, albeit at noise levels and frequencies well beyond the original goals of LISA Pathfinder and even below the LISA requirements. Several remaining noise sources cannot be effectively quantified by dedicated on-board measurements and cannot be excluded on the basis of our ground testing to date. Forces acting on the TM surface include other spontaneous outgassing and/or virtual leak pressure effects, from the many cables and other surfaces inside the GRS, as well as electrostatic noise from fluctuating small-scale surface patch potentials interacting with their own steady field values. Regarding the electrostatic noise, while the interaction of the spatially averaged stray fields and the TM charge is well measured in flight [15], random field fluctuations on smaller scales cannot be excluded, at the levels of the remaining noise, from either flight measurements or ground studies [16–18]. Pressure-related effects can be addressed by direct pressure measurements in representative GRS hardware, while electrostatic force noise and other short range forces can be studied on the ground in torsion pendulum tests with much smaller gaps between the TM and surrounding conducting surfaces.

Several force noise sources coupling to the bulk of the TM can also not be excluded to date. These include relatively high-frequency (> 10 Hz) magnetic fields "down-converting" into the LPF band by the $\nabla(B^2)$ dependence of the magnetic force. These could be tested in the future by audio-range magnetometers, which were not available on board LPF. Gravitation from various elements of the satellite is also a relevant low-frequency noise source via several mechanisms. The depletion of N_2 from the cold gas micropropulsion system produces a deterministic gravitational drift in Δg, typically several hundreds of fm s^{-2}/day during LPF operations. Pressure fluctuations in the thruster plumbing and flow irregularities may cause gravitational noise at a relevant level. Gravity fluctuations from outgassing of the payload enclosure and spacecraft could also be relevant, considering that mass loss after years in orbit can amount to hundreds of grams per year, as observed in other deep space missions like Rosetta [19]. The stochastic portion of this mass loss could also contribute to a f^{-2} gravitational PSD. These self-gravity effects can be addressed on the ground by dedicated pressure and vacuum tests with model hardware.

Before closing the discussion on the low-frequency LPF noise, we note that the 7 month period presented in Fig. 2—referred to as group B—falls between two other shorter periods, groups A and C, for which we have two runs each, totalling roughly 15 days of measurement. Both groups are noisier than group B (see Fig. 2). Though the differences in noise between group B and groups A and C are not yet understood, some aspects of the different conditions for the different groups, summarized in Table 1 of Ref. [10], merit discussion. The first group, measured in March–April 2016 and presented in Ref. [3], was performed at significantly higher pressures and outgassing rates. The measured

temperature dependence of Δg, visible both in dedicated temperature modulation experiments and in correlations in the noise data, was roughly 4 $\mathrm{pm/s^{-2}\ K^{-1}}$ in group A and significantly lower in the later groups; while a correction for this effect made a visible impact only at 20 μHz and lower frequencies, we cannot exclude that other pressure-related effects were relevant to the early data. Additionally, groups A and C were performed using different tanks of N_2 for the cold gas micropropulsion, with different hardware in different positions and thus with different levels of the gravitation noise contribution. Finally, we note that group C was measured in the final few weeks of the mission, just after a final programed cooling of the payload to near 0 °C that resulted in a dramatic increase in the frequency of glitches, to be discussed shortly. We aborted the run by raising the temperature to roughly 22 °C, apparently recovering the normal glitch rate, but, as said, the differential acceleration noise at low frequencies could not recover the same performance measured in group B.

The apparent force glitches mentioned earlier merit further discussion. Although they can be removed from the data as discussed above, their presence would increase the ambiguity in the signal-dominated LISA data, thus also increasing the difficulty of data analysis and effectively limiting the observatory sensitivity to unexpected transient GW signals. The nature and potential sources of the observed glitches, and possible strategies to limit their number, are key issues for LISA.

The physical origin of the glitches remains unknown. A more detailed discussion of the investigation into their nature can be found in Ref. [20], and we simply summarize the conclusions here. The glitches produce no permanent change in Δg and have no counterpart events in the differential torque on the TM, in the various TM and spacecraft dynamical degrees of freedom, or in any of the environmental and instrument monitors on board. These observations tend to exclude a number of otherwise suspect physical mechanisms, including mechanical relaxation on board LPF, spurious voltages on the GRS electrodes from its electronics, micrometeorite impacts, spikes in spacecraft or TM control actuators, laser intensity pulses, and magnetic or radiation anomalies. A large increase in the glitch rate, to roughly 30 events per day, was observed upon decreasing the LPF payload temperature to near 0 °C [14]. The glitch rate was then observed to slowly decrease over one week of observation at this low temperature. The glitch rate decreased back to its original rate, below 1/day, upon restoring the temperature to roughly 10 °C. This observed change with temperature could suggest as the cause of the glitches thermomechanically liberated gas "bursts" inside the GRS vacuum chambers housing the TM, from, for instance, cables, caging piezomotors, and various mechanical interfaces including the gravitational balance masses [9]. The number of molecules needed to explain the observed impulses correspond to pressure increases of the order of nPa, on top of roughly μPa background pressure. Testing on the ground for such pressure increases appears possible, as would impulse detection in dedicated torsion pendulum tests in a representative geometry with the relevant hardware [21]. Appropriate testing thus looks possible for the glitch question, which will demand further study and experimentation in moving from LPF to LISA.

Conclusions: Implications for LISA gravitational-wave observation.—The LPF differential acceleration measurement noise, presented in Fig. 1, represents a striking demonstration of the ability to use an LPF-like geodesic reference system—with a roughly 2 kg test mass free-falling, contact-free, inside a drag-free spacecraft and tracked by an interferometric displacement readout—with the precision needed to do gravitational-wave science from space at frequencies as low as 20 μHz. The performance is better than the proposed LISA requirement at all frequencies, and its achievement in a 2.5 million km triangular orbiting constellation, with the required interspacecraft laser interferometry, will allow, with a margin, the full observation program of the LISA observatory recently proposed and selected by ESA [2].

While a full discussion of the observational impacts of this result is deferred to more dedicated articles, we cite a few examples of how LPF performance is a critical, wide-band experimental benchmark for gravitational-wave astrophysics. For the example of the coalescence of two supermassive black holes (SMBHs), each with 5×10^6 solar mass at a redshift $z = 2$, the dominant harmonic of the signal will enter the band at 20 μHz, 35 days before coalescence, accumulating a SNR of 50 by the time the source crosses 100 μHz, just 10 h before the final merger at roughly 0.8 mHz (with a final total SNR in excess of 1400). Extending the observation into the 20–100 μHz band thus increases the observation time by nearly 2 orders of magnitude for this source, thus enabling a premerger detection with a sky localization better than 100 square degrees in 40% of the cases. This significantly improves the possibility to raise alerts to other observatories to detect electromagnetic signatures of the final merger. At millihertz frequencies, we note that the improvement in the Brownian white noise floor, to roughly 1.3 $\mathrm{fm\,s^{-2}/\sqrt{Hz}}$ for a single TM, will allow the resolution of more galactic compact object binaries but also a more reliable measurement of the "foreground" noise of such galactic sources. The entire signal power of the cited cosmic SMBHs and the local galactic foreground fall in the band below 4 mHz where the LISA target sensitivity is expected to be dominated by TM acceleration noise and for which the LPF differential acceleration results presented here thus represent an important experimental verification.

This work has been made possible by the LISA Pathfinder mission, which is part of the space-science program of the European Space Agency. The French

contribution has been supported by the CNES (Accord Specific de projet CNES 1316634/CNRS 103747), the CNRS, the Observatoire de Paris, and the University Paris-Diderot. E. P. and H. I. also acknowledge the financial support of the UnivEarthS Labex program at Sorbonne Paris Cité (ANR-10-LABX-0023 and ANR-11-IDEX-0005-02). The Albert-Einstein-Institut acknowledges the support of the German Space Agency, DLR. The work is supported by the Federal Ministry for Economic Affairs and Energy based on a resolution of the German Bundestag (FKZ 50OQ0501 and FKZ 50OQ1601). The Italian contribution has been supported by Agenzia Spaziale Italiana and Instituto Nazionale di Fisica Nucleare. The Spanish contribution has been supported by Contracts No. AYA2010-15709 (MICINN), No. ESP2013-47637-P, and No. ESP2015-67234-P (MINECO). M. N. acknowledges support from Fundacion General CSIC (Programa ComFuturo). F. R. acknowledges an FPI contract (MINECO). The Swiss contribution acknowledges the support of the Swiss Space Office (SSO) via the PRODEX Program of ESA. L. F. acknowledges the support of the Swiss National Science Foundation. The United Kingdom groups acknowledge support from the United Kingdom Space Agency (UKSA), the University of Glasgow, the University of Birmingham, Imperial College, and the Scottish Universities Physics Alliance (SUPA). J. I. T. and J. S. acknowledge the support of the U.S. National Aeronautics and Space Administration (NASA). N. Korsakova acknowledges the support of the Newton International Fellowship from the Royal Society.

[*] Deceased.

[1] P. McNamara, S. Vitale, and K. Danzmann, Classical Quantum Gravity **25**, 114034 (2008).
[2] P. Amaro-Seoane *et al.*, arXiv:1702.00786.
[3] M. Armano *et al.*, Phys. Rev. Lett. **116**, 231101 (2016).
[4] S. Anza *et al.*, Classical Quantum Gravity **22**, S125 (2005).
[5] F. Antonucci *et al.*, Classical Quantum Gravity **28**, 094002 (2011).
[6] G. Heinzel, C. Braxmaier, R. Schilling, A. Rüdiger, D. Robertson, M. te Plate, V. Wand, K. Arai, U. Johann, and K. Danzmann, Classical Quantum Gravity **20**, S153 (2003).
[7] G. Heinzel, V. Wand, A. García, O. Jennrich, C. Braxmaier, D. Robertson, K. Middleton, D. Hoyland, A. Rüdiger, R. Schilling, U. Johann, and K. Danzmann, Classical Quantum Gravity **21**, S581 (2004).
[8] H. Audley *et al.*, Classical Quantum Gravity **28**, 094003 (2011).
[9] M. Armano *et al.*, Classical Quantum Gravity **33**, 235015 (2016).
[10] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.120.061101 for a section on glitch removal.
[11] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.120.061101 for Table 1.
[12] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.120.061101 for a section on the Bayesian ASD estimate.
[13] A. Cavalleri, G. Ciani, R. Dolesi, A. Heptonstall, M. Hueller, D. Nicolodi, S. Rowan, D. Tombolato, S. Vitale, P. J. Wass, and W. J. Weber, Phys. Rev. Lett. **103**, 140601 (2009).
[14] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.120.061101 for a section on the estimate of the excess over Brownian noise.
[15] M. Armano *et al.*, Phys. Rev. Lett. **118**, 171101 (2017).
[16] F. Antonucci, A. Cavalleri, R. Dolesi, M. Hueller, D. Nicolodi, H. B. Tu, S. Vitale, and W. J. Weber, Phys. Rev. Lett. **108**, 181101 (2012).
[17] S. E. Pollack, S. Schlamminger, and J. H. Gundlach, Phys. Rev. Lett. **101**, 071101 (2008).
[18] H. Yin, Y. Z. Bai, M. Hu, L. Liu, J. Luo, D. Y. Tan, H. C. Yeh, and Z. B. Zhou, Phys. Rev. D **90**, 122001 (2014).
[19] B. Schläppi, K. Altwegg, H. Balsiger, M. Hässig, A. Jäckel, P. Wurz, B. Fiethe, M. Rubin, S. A. Fuselier, J. J. Berthelier, J. De Keyser, H. Rème, and U. Mall, J. Geophys. Res. **115**, A12313 (2010).
[20] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.120.061101 for a section on glitch properties.
[21] A. Cavalleri, G. Ciani, R. Dolesi, M. Hueller, D. Nicolodi, D. Tombolato, P. J. Wass, W. J. Weber, S. Vitale, and L. Carbone, Classical Quantum Gravity **26**, 094012 (2009).