SUPPLEMENTAL MATERIAL

Limits on Environmental Noise Couplings

We have carried out the same set of tests for systematic noise biases as outlined in §6.4 of Ref. [1] for the first-generation Holometer experiment, with some additions and extensions as described here. The results of these tests show there would be insignificant impact on our science results from environmental RF noise. The measured limits on environmental noise couplings are shown in Fig. 3.

Input Laser Noise

In the original straight-arm configuration, the search for an exotic signal was carried out at frequencies below 10 MHz. Consequently, characterization efforts for correlated laser noise targeted the band below 10 MHz. For the rotational configuration, to account for the possibility of a signal extending to 20 MHz, the measured limits on correlated laser noise must be extended comensurately.

Laser Phase Noise. Measurements of the interferometer transfer function of laser phase noise to DARM were carried out as described in §6.4.2 of Ref. [1], but with the measurement bandwidth extended to 20 MHz. The measurements were made on several nonconsecutive days and the results compared for consistency. Above 10 MHz, the phase of the transfer function was found to vary from day to day. This is likely caused by alignment-dependent variation of the couplings of higher-order spatial modes (HOMs) to the antisymmetric-port (ASP) photodetectors (PDs), as occurs upon each re-tuning of the instrument. Because the interferometers do not have output mode cleaner cavities, HOMs are not attenuated prior to detection. HOM noise sidebands beat with the fundamental-mode carrier to produce amplitude modulations of the detected power. To obtain a worst-case limit shown in Fig. 3 (orange curve), we have taken the maximum coupling amplitude among all measurements for each frequency bin individually.

Laser Amplitude Noise. Measurements of the interferometer transfer function of laser amplitude noise to DARM were carried out as described in §6.4.1 of Ref. [1], but with the measurement bandwidth extended to 20 MHz. They were found to be consistent over several days. However, because this measurement passively relies on the amplitude noise of the input laser, the coherence is poor above 10 MHz, where deviations in the phase noise case were most strongly observed. Fig. 3 (blue curve) shows the amplitude noise coupling obtained from the transfer function measurement average.

Ambient Radiofrequency Noise

Ambient RF coupling to the interferometer signal takes two possible forms. The first is direct coupling to the post-detection electronics. Tests of this form of contamination (see §6.4.3 of [1]) were extended to include a new white noise test at full ASP PD photocurrent. The second is excitation of the actual DARM signal, either by coupling to physical motion—e.g., exciting the piezoelectric transducer (PZT) actuators of the end-mirrors—or by an optically active effect in the optical substrate or reflective coatings, coupling electric or magnetic fields to path length. The actuations of the PZTs were tested by coupling electrical signals, either from the ambient RF environment or capacitively, to the PZT signal cables directly. These effects have not been newly tested since the first configuration. However, the addition of a 45°-incidence bend mirror in one arm of each interferometer introduces potential new modes of coupling directly to DARM via its sensitivity to polarization. Tests for polarization-dependent effects are detailed further below.

Post-Detection Electronics Noise. As in the previous configuration, the most sensitive test of correlated post-detection additive electronic noise is a run with the photodiodes, amplifiers, and analog-to-digital converter (ADC) systems operating with the laser light off. Because the system under normal operation is highly shot-noise-limited, a relatively short 10 hour dark run serves to test the possibility of additive noise coupling to both detector signal chains and causing an apparent correlated signal. This test shows no evidence of this source of correlated signal. However, the dark test leaves open the possibility that high-noise, high-current operating conditions nonlinearly impart correlations not apparent under dark conditions. Large noise on an ADC having digitization issues, an amplifier or PD with saturation problems, or amplifier nonlinearity such as slew rate limitations could couple the detection system to external signals differently during data runs than during dark noise tests.

For this reason, the system was newly run with a separate halogen light bulb focused on each of the ASP PDs so that the photocurrent was the same as during science-mode operation. The light bulbs were optically coupled to the photodiodes with fast lenses and run with the lowest possible light bulb currents to avoid UV illumination on the photodiodes during the test. As during the science data runs, the photocurrent to shot noise ratio was monitored to verify the noise purity. The disadvantage over the dark noise test is that bright noise data must be collected for as long as the science data integration to reach the needed sensitivity to spurious correlation. We obtained 745 hours of data in this configuration, which is less than the total integration time for the science-mode configuration. Hence the statistics-limited electronics noise coupling, shown in Fig. 3 (green

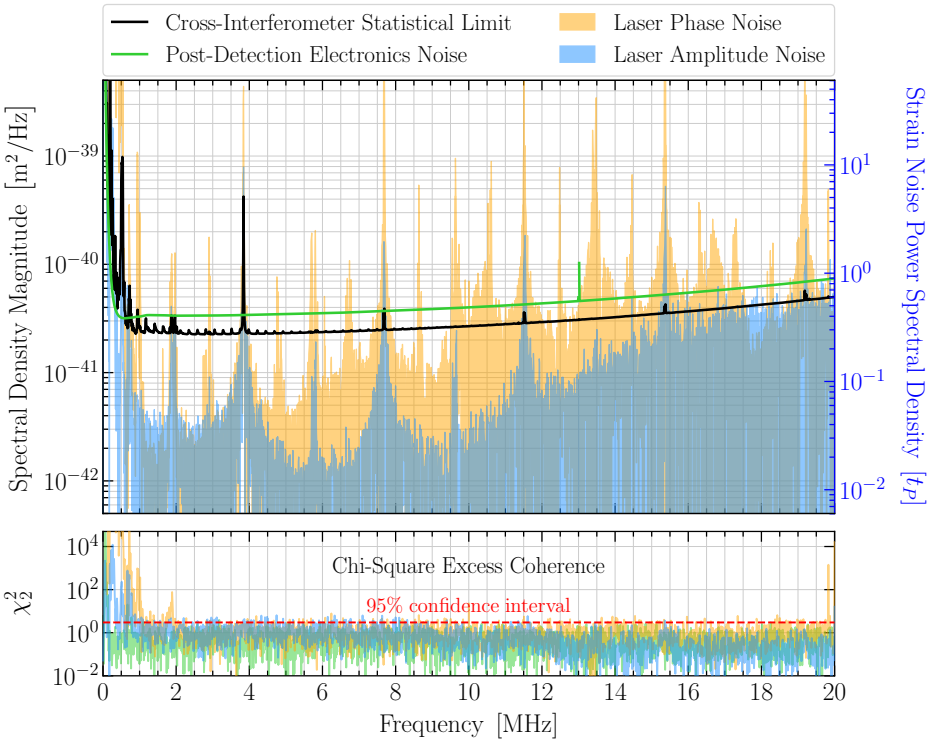

FIG. 3. Measured limits on environmental sources of correlated noise. The curves show the exclusions of input laser phase noise (orange), input laser amplitude noise (blue), and post-detection electronics noise (green). For reference, the black curve shows the statistical sensitivity of the two-interferometer cross-spectrum. The electronics noise curve is from a 745 hour ex-situ measurement, in contrast with the in-situ sensitivity of the 1098 hour science data set. The reduced χ^2 statistic for $k = 2$ degrees of freedom, shown in the bottom panel, indicates the level of statistical certainty of coherent noise detection. Values > 3 are indicative of significance at the 95% confidence level.

curve), sets a slightly worse limit than the sensitivity of the science data itself (black curve). However, no correlation was found in these data over the range from 1.1 MHz to 20 MHz; this limit could be improved with a longer integration time.

Polarization Noise. New polarization-dependent couplings are introduced due to the 45° angle of incidence of the beam at the bend mirrors. Both bend mirrors are coated for p-polarization. However, any small rotation of the polarization axis introduces some s-polarization at the bend mirror surface, modulating the penetration depth into the dielectric coating layers. Through this mechanism, ambient magnetic fields rotating the polarization of the light could modulate both the amplitude and phase of the reflected laser field. Such effects are negligible at the power recycling mirror and end mirrors, which have near-normal angles of incidence, and cancel to first order at the 45°-incidence beamsplitter, due to the double-passing of the light.

To check for such a possibility, a series of RF electromagnetic tests were conducted on the bend mirrors while the interferometers were running. A set of Helmholz drive coils were placed around both 10 inch stainless steel corner cubes containing the bend mirrors and energized with a power amplifier operating to 15 MHz. The coils were placed so that they were aligned along three perpendicular axes. A pickup coil placed between the cubes measured both the environmental RF and the drive coil signal. A comprehensive set of measurements were first made relative to ambient RF fields. Then, during the drive tests, a second pickup antenna was used as a witness sensor to maximize the sensitivity of the transfer function upper limit between the magnetic field at the pickup coil and DARM, similar to the method in Eq. 25 of Ref. [1]. The pickup coil and witness signals were digitized at 50 MHz together with the interferometer signals and cross-correlated. Several hours of RF noise with magnetic field amplitudes 1000 times ambient was measured, using swept RF signals, broadband noise, and several discrete lines. However, the measurement was found to be limited by direct couplings of the detection electronics to the strong RF field being generated, as the measured transfer function was identical even when the interferometers were not locked. This indicates that the strongest couplings of the apparatus to ambient electromagnetic fields are not optical but electronic in nature. Thus, if ambient electromagnetic fields of sufficient strength to induce an optical effect were present, they would have also been detected as correlated post-detection electronic noise.

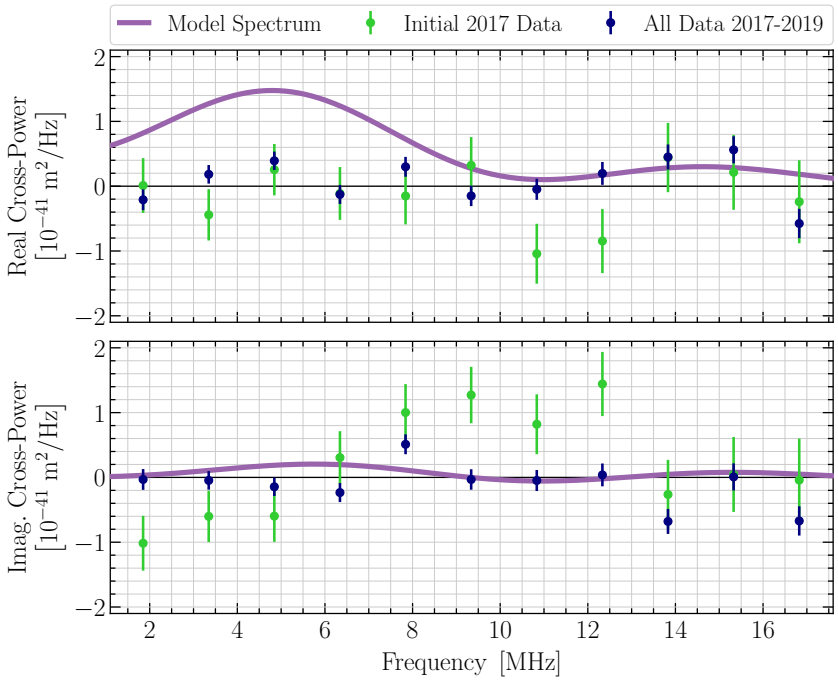

FIG. 4. Interferometer CSD from Spring 2017 data (125 hours; green data points) vs. the full data set (1098 hours; blue data points). The data are shown for both real and imaginary components at 1.5 MHz resolution. For reference, all panels are overlaid with a semiclassical model spectrum of quantum geometrical fluctuations (purple curves) reproduced from Ref. [6], but adjusted to account for the actual dimensions of the as-built instrument.

Chronology of Observing Runs

The measurements reported here were taken over a two-year period, from April 2017 to August 2019, in five separate data runs. Almost all data were taken at night for more stable optical cavity lock, as the reconfigured bent-arm instrument had less insulation and could not maintain long-term alignment when direct sunlight caused varying thermal expansion in the vacuum tubes.

In Spring 2017, the first 125 hours of data were collected (run 1), passing all systematics tests established in the first-generation Holometer experiment prior to the reconfiguration [1]. The data showed an apparent broadband correlation from 6.5 MHz to 13 MHz purely in the *imaginary* part of the cross spectrum, as shown in Fig. 4 (green data points). Since this feature was entirely different from the predicted model, its statistical significance was difficult to assess post facto in a model-independent or nonparametric manner.

To further investigate this effect, we collected two more sets of data throughout Summer and Fall 2017: one for 113 hours of additional science data in the nominal configuration (run 2), and another for 141 hours of data in an "inverted fringe" configuration (run 3). In the latter, one of the control systems was operated with a sign inversion of the DARM offset and feedback gain, so that one interferometer was locked on the opposite side of the dark fringe (IFO 1 inverted for 100 hours, IFO 2 for 41 hours; see Figs. 1 and 2). This has the effect of inverting the sign of the optical response to DARM perturbations, as well as to other sources of phase noise, creating a phase inversion in the interferometer cross spectrum. The response to amplitude noise sources, including post-detection additive electronics noise, is unchanged. This technique is thus a diagnostic tool for determining the coupling mechanism of potential background cross-correlations. Our tests did not clearly reproduce the Spring 2017 signal, however, and further data acquisition was precluded by the rapidly decaying power levels of the two Mephisto 2 W Nd:YAG lasers (both after over 30,000 hours of use).

In Spring 2018, the lasers were refurbished with replacement pump diodes by Coherent, Inc. and reinstalled into the interferometers. This restored the laser power to levels even higher than before. In Summer 2018, only a limited set of science data amounting to 64 hours was collected (run 4), due to interferometer controls difficulties. This had insufficient statistical power to independently test the first data set. Additional systematic tests were conducted throughout the remainder of 2018, in parallel with efforts to retune the control systems for higher operational stability. From Spring through Summer 2019, a final, extended science run was conducted with reoptimized control systems, yielding 654 hours of additional data (run 5). With the combined statistical power of all data sets since Spring 2017, we obtained a conclusive rejection of the apparent effect in the first data run.