Searching for stochastic gravitational waves using data from the two co-located LIGO Hanford detectors

Canadian Institute for Theoretical Astrophysics, University of Toronto, Toronto, Ontario, M5S 3H8, Canada
Tsinghua University, Beijing 100084, China
University of Michigan, Ann Arbor, MI 48109, USA
Rochester Institute of Technology, Rochester, NY 14623, USA
INFN, Sezione di Roma Tor Vergata, I-00133 Roma, Italy
National Tsing Hua University, Hsinchu Taiwan 300
Charles Sturt University, Wagga Wagga, NSW 2678, Australia
Caltech-CaRT, Pasadena, CA 91125, USA
Pusan National University, Busan 609-735, Korea
Australian National University, Canberra, ACT 0200, Australia
Carleton College, Northfield, MN 55057, USA
University of Roma Tor Vergata, I-00133 Roma, Italy
Università di Roma ‘La Sapienza’, I-00185 Roma, Italy
The George Washington University, Washington, DC 20052, USA
University of Cambridge, Cambridge, CB2 1TN, United Kingdom
University of Minnesota, Minneapolis, MN 55455, USA
The University of Sheffield, Sheffield S10 2TN, United Kingdom
Wigner RCP, RMKI, H-1121 Budapest, Konkoly Thege Miklós ut 29-33, Hungary
University of Sannio at Benevento, I-84100 Benevento, Italy and INFN (Sezione di Napoli), Italy
Inter-University Centre for Astronomy and Astrophysics, Pune - 411007, India
Università di Trento, I-38123 Povo, Trento, Italy
INFN, Gruppo Collegato di Trento, I-38050 Povo, Trento, Italy
California Institute of Technology, Pasadena, CA 91125, USA
Northwestern University, Evanston, IL 60208, USA
Montclair State University, Montclair, NJ 07043, USA
The Pennsylvania State University, University Park, PA 16802, USA
MTA Eötvös University, ‘Lendulet’ Astrophysics Research Group, Budapest 1117, Hungary
National Astronomical Observatory of Japan, Tokyo 181-8588, Japan
Università di Perugia, I-06123 Perugia, Italy
University of Szeged, Dóm tér 9, Szeged 6720, Hungary
Rutherford Appleton Laboratory, HSIC, Chilton, Didcot, Oxon, OX11 0QX, United Kingdom
Embry-Riddle Aeronautical University, Prescott, AZ 86301, USA
Perimeter Institute for Theoretical Physics, Ontario, N2L 2Y5, Canada
American University, Washington, DC 20016, USA
College of William and Mary, Williamsburg, VA 23187, USA
University of Adelaide, Adelaide, SA 5005, Australia
Raman Research Institute, Bangalore, Karnataka 560080, India
Korea Institute of Science and Technology Information, Daejeon 305-806, Korea
University of Białystok, 15-424 Białystok, Poland
University of Southampton, Southampton, SO17 1BJ, United Kingdom
IISER-TVM, CET Campus, Trivandrum Kerala 695016, India
Hobart and William Smith Colleges, Geneva, NY 14456, USA
Institute of Applied Physics, Nizhny Novgorod, 603950, Russia
Seoul National University, Seoul 151-742, Korea
Hanyang University, Seoul 133-791, Korea
NCBJ, 05-400 ´Swierk-Otwock, Poland
IM-PAN, 00-956 Warsaw, Poland
Institute for Plasma Research, Bhat, Gandhinagar 382428, India
Utah State University, Logan, UT 84322, USA
The University of Melbourne, Parkville, VIC 3010, Australia
University of Brussels, Brussels 1050 Belgium
SUPA, University of Strathclyde, Glasgow, G1 1XQ, United Kingdom
ESPCI, CNRS, F-75005 Paris, France
Università di Camerino, Dipartimento di Fisica, I-62032 Camerino, Italy
The University of Texas at Austin, Austin, TX 78712, USA
Southern University and A&M College, Baton Rouge, LA 70813, USA
IISER-Kolkata, Mohanpur, West Bengal 741252, India
National Institute for Mathematical Sciences, Daejeon 305-390, Korea
RRCAT, Indore MP 452013, India
Tata Institute for Fundamental Research, Mumbai 400005, India
Louisiana Tech University, Ruston, LA 71272, USA
SUPA, University of the West of Scotland, Paisley, PA1 2BE, United Kingdom
Institute of Astronomy, 65-265 Zielona Góra, Poland
Indian Institute of Technology, Gandhinagar Ahmedabad Gujarat 382424, India
I. INTRODUCTION

The detection of a stochastic gravitational-wave background (SGWB) of either cosmological or astrophysical origin, is a major science goal for both current and planned searches for gravitational waves (GWs) [1–4]. Given the weakness of the gravitational interaction, cosmological GWs are expected to decouple from matter in the early universe much earlier than any other form of radiation (e.g., photons, neutrinos, etc.). The detection of such a primordial GW background by the current ground-based detectors [5, 6], proposed space-based detectors [7, 8], or a pulsar timing array [9, 10] would give us a picture of the universe mere fractions of a second after the Big-Bang [1–3, 12], allowing us to study the physics of the highest energy scales, unachievable in standard laboratory experiments [4]. The recent results from the BICEP2 experiment indicate the existence of cosmic microwave background B-mode polarization at degree angular scales [13], which may be due to an ultra-low frequency primordial GW background, such as would be generated by amplification of vacuum fluctuations during cosmological inflation; however, it cannot currently be ruled out that the observed B-mode polarization is due to a Galactic dust foreground [14, 15]. These GWs and their high frequency counterparts in standard slow-laser experiments [3, 4] are at a sufficiently low level that they are not observable in coherence measurements over the sensitivity levels of current and advanced LIGO detectors. Hence, they are not the target of our current analysis. However, many non-standard inflationary models predict GWs that could be detected by advanced LIGO detectors.

On the other hand, the detection of a SGWB due to spatially and temporally unresolved foreground astrophysical sources such as magnetars [16], rotating neutron stars [17], galactic and extragalactic compact binaries [18, 20], or the inspiral and collisions of supermassive black holes associated with distant galaxy mergers [21], would provide information about the spatial distribution and formation rate of these various source populations.

Given the random nature of a SGWB, searches require cross-correlating data from two or more detectors [1, 22–25], under the assumption that correlated noise between any two detectors is negligible. For such a case, the contribution to the cross-correlation from the (common) GW signal grows linearly with the observation time $T$, while that from the noise grows like $\sqrt{T}$. Thus, the signal-to-noise ratio (SNR) also grows like $\sqrt{T}$. This allows one to search for stochastic signals buried within the detector noise by integrating for a sufficiently long interval of time.

For the widely-separated detectors in Livingston, LA and Hanford, WA, the physical separation ($\sim 3000$ km) eliminates the coupling of local instrumental and environmental noise between the two detectors, while global disturbances such as electromagnetic resonances are at a sufficiently low level that they are not observable in coherence measurements between the (first-generation) detectors at their design sensitivity [3, 24, 30].

While physically-separated detectors have the advantage of reduced correlated noise, they have the disadvantage of reduced sensitivity to a SGWB; physically-separated detectors respond at different times to GWs from different directions and with differing response amplitudes depending on the relative orientation and (mis)alignment of the detectors [23–25]. Co-located and co-aligned detectors, on the other hand, such as the 4 km and 2 km interferometers in Hanford, WA (denoted H1 and H2), respond identically to GWs from all directions and for all frequencies below a few kHz. They are thus, potentially, an order-of-magnitude more sensitive to a
we summarize our results and discuss potential improvements. We describe the cross-correlation procedure used to search for correlated instrumental and environmental noise, given that the two detectors share the same local environment. Methods to identify and mitigate the effects of correlated noise are thus needed to realize the potential increase in sensitivity of co-located detectors.

In this paper, we apply several noise identification and mitigation techniques to data taken by the two LIGO Hanford detectors, H1 and H2, during LIGO’s fifth science run (S5, November 4, 2005, to September 30, 2007) in the context of a search for a SGWB. This is the first stochastic analysis using LIGO science data that addresses the complications introduced by correlated environmental noise. As discussed in the references [29, 30], the coupling of global magnetic fields to non-colocated advanced LIGO detectors could produce significant correlations between them thereby reducing their sensitivity to SGWB by an order of magnitude. We expect the current H1-H2 analysis to provide a useful precedent for SGWB searches with advanced detectors in such (expected) correlated noise environment.

Results are presented at different stages of cleaning applied to the data. We split the analysis into two parts—one for the frequency band 460–1000 Hz, where we are able to successfully identify and exclude significant narrow-band correlations; and the other for the band 80–160 Hz, where even after applying the noise reduction methods there is still evidence of residual contamination, resulting in a large systematic uncertainty for this band. The frequencies below 80 Hz and between 160–460 Hz are not included in the analysis because of poor detector sensitivity and contamination by known noise artifacts. We observe no evidence of a SGWB and so our final results are given in the form of upper-limits. Due to the presence of residual correlated noise between 80–160 Hz, we do not set any upper-limit for this frequency band. Since we do not observe any such residual noise between 460–1000 Hz, in that frequency band and the 5 sub-bands assigned to it, we set astrophysical upper-limits on the energy density of stochastic GWs.

The rest of the paper is organized as follows. In Sec. II we describe sources of correlated noise in H1 and H2, and the environmental and instrumental monitoring system. In Sec. III we describe the cross-correlation procedure used to search for a SGWB. In Secs. IV and V we describe the methods that we used to identify correlated noise, and the steps that we took to mitigate it. In Secs. VI and VII we give the results of our analysis applied to the S5 H1-H2 data. Finally, in Sec. VIII we summarize our results and discuss potential improvements to the methods discussed in this paper.

II. COMMON NOISE IN THE TWO LIGO HANFORD DETECTORS

At each of the LIGO observatory sites the detectors are supplemented with a set of sensors to monitor the local environment [5, 31]. Seismometers and accelerometers measure vibrations of the ground and various detector components; microphones monitor acoustic noise; magnetometers monitor magnetic fields that could couple to the test masses (end mirrors of the interferometers) via the magnets attached to the test masses to control their positions; radio receivers monitor radio frequency (RF) power around the laser modulation frequencies, and voltage line monitors record fluctuations in the AC power. These physical environment monitoring (PEM) channels are used to detect instrumental and environmental disturbances that can couple to the GW strain channel. We assume that these channels are completely insensitive to GW strain. The PEM channels are placed at strategic locations around the observatory, especially near the corner and ends of the L-shaped interferometer where important laser, optical, and suspension systems reside in addition to the test masses themselves.

Information provided by the PEM channels is used in many different ways. The most basic application is the creation of numerous data quality flags identifying stretches of data that are corrupted by instrumental or environmental noise [32]. The signals from PEM channels are critical in defining these flags: microphones register airplanes flying overhead, seismometers and accelerometers detect elevated seismic activity or anthropogenic events (trucks, trains, logging), and magnetometers detect fluctuations in the mains power supply and the Earth’s magnetic field.

In searches for transient GW signals, such as burst or coalescing binary events, information from the PEM channels has been used to construct vetoes [33, 50]. When a clear association can be made between a measured environmental event and a coincident glitch in the output channel of the detector, then these times are excluded from the transient GW searches. These event-by-event vetoes exclude times of order hundreds of milliseconds to a few seconds.

Similarly, noise at specific frequencies, called noise lines, can affect searches for GWs from rotating neutron stars or even for a SGWB. In S5, data from PEM channels were used to verify that some of the apparent periodic signals were in fact due to noise sources at the observatories [37, 38]. Typically the neutron-star search algorithms can also be applied to the PEM data to find channels that have noise lines at the same frequencies as those in the detector output channel. The coherence is also calculated between the detector output and the PEM channels, and these results provide additional information for determining the source of noise lines.

The study of noise lines has also benefited past LIGO searches for stochastic GWs. For example, in LIGO’s search for a SGWB using the data from the S4 run [27], correlated noise between the Hanford and Livingston detectors was observed in the form of a forest of sharp 1 Hz harmonic lines. It was subsequently determined that these lines were caused by the sharp ramp of a one-pulse-per-second signal, injected into the data acquisition system to synchronize it with the Global Positioning System (GPS) time reference. In the S5 stochastic search [28], there were other prominent noise lines that were subsequently identified through the use of the PEM signals.

In addition to passive studies, where the PEM signals are observed and associations are made to detector noise, there have also been a series of active investigations where noise
was injected into the detector environment in order to measure its coupling to the GW channel. Acoustic, seismic, magnetic, and RF electromagnetic noise were injected into the observatory environment at various locations and responses of the detectors were studied. These tests provided clues and ways to better isolate the detectors from the environment.

All the previous LIGO searches for a SGWB have used the physically-separated Hanford and Livingston detectors and assumed that common noise between these non-colocated detectors was inconsequential. This assumption was strongly supported by observations—i.e., none of the coherence measurements performed to date between these detectors revealed the presence of correlations other than those known to be introduced by the instrument itself (for example, harmonics of the 60 Hz power line). Since the analysis presented here uses the two co-located Hanford detectors, which are susceptible to correlated noise due to the local environment, new methods were required to identify and mitigate the correlated noise.

III. CROSS-CORRELATION PROCEDURE

The energy density spectrum of SGWB is defined as

$$\Omega_{gw}(f) \equiv \frac{f}{\rho_c} \frac{\rho_{gw}}{df}$$

where \(\rho_c = \frac{3c^2H_0^2}{8\pi G}\) is the critical energy density and \(\rho_{gw}\) is the GW energy density contained in the frequency range \(f\) and \(f + df\). Since most theoretical models of stochastic backgrounds in the LIGO band are characterized by a power-law spectrum, we will assume that the fractional energy density in GWs \(^{[39]}\) has the form

$$\Omega_{gw}(f) = \Omega_\alpha \left(\frac{f}{f_{\text{ref}}}\right)^\alpha,$$

where \(\alpha\) is the spectral index and \(f_{\text{ref}}\) is some reference frequency. We will consider two values for the spectral index: \(\alpha = 0\) which is representative of many cosmological models, and \(\alpha = 3\) which is characteristic of many astrophysical models. This latter case corresponds to a flat (i.e., constant) one-sided power spectral density (PSD) in the strain output of a detector \(S_{gw}(f)\), since

$$S_{gw}(f) = \frac{3H_0^2}{10\pi^2} \frac{\Omega_{gw}(f)}{f^3} \propto f^{\alpha-3}.$$ 

Here \(H_0\) is the present value of the Hubble parameter, assumed to be \(H_0 = 68\) km/s/Mpc \(^{[40]}\).

Following the procedures described in \(^{[23]}\), we construct our cross-correlation statistic as estimators of \(\Omega_\alpha\) for individual frequency bins, of width \(\Delta f\), centered at each (positive) frequency \(f\). These estimators are simply the measured values of the cross-spectrum of the strain output of two detectors divided by the expected shape of the cross-correlation due to a GW background with spectral index \(\alpha\):

$$\hat{\Omega}_\alpha(f) \equiv \frac{2}{T} \frac{\Re [\tilde{s}_1(f)\tilde{s}_2(f)]}{\gamma(f)S_\alpha(f)}.$$

Here \(T\) is the duration of the data segments used for Fourier transforms; \(\tilde{s}_1(f), \tilde{s}_2(f)\) are the Fourier transforms of the strain time-series in the two detectors; \(S_\alpha(f)\) is proportional to the assumed spectral shape,

$$S_\alpha(f) \equiv \frac{3H_0^2}{10\pi^2} \frac{1}{f^3} \left(\frac{f}{f_{\text{ref}}}\right)^\alpha;$$

and \(\gamma(f)\) is the overlap reduction function \(^{[22,23]}\), which encodes the reduction in sensitivity due to the separation and relative alignment of the two detectors. For the H1-H2 detector pair, \(\gamma(f) \approx 1\) for all frequencies below a few kHz \(^{[41]}\).

In the absence of correlated noise, one can show that the above estimators are optimal—i.e., they are unbiased, minimal-variance estimators of \(\Omega_\alpha\) for stochastic background signals with spectral index \(\alpha\). Assuming that the detector noise is Gaussian, stationary, and much larger in magnitude than the GW signal, the expectation value of the variance of the estimators is given by

$$\sigma_{\hat{\Omega}_\alpha}(f) \approx \frac{1}{2T\Delta f} \frac{P_1(f)P_2(f)}{\gamma^2(f)S_\alpha^2(f)},$$

where \(P_1(f), P_2(f)\) are the one-sided PSDs of the detector output \(\tilde{s}_1(f), \tilde{s}_2(f)\) respectively. For a frequency band consisting of several bins of width \(\Delta f\), the optimal estimator and corresponding variance are given by the weighted sum

$$\hat{\Omega}_\alpha = \sum_f \frac{\sigma_{\hat{\Omega}_\alpha(f)}^2}{\sigma_{\alpha(f)}^2}, \quad \sigma_{\hat{\Omega}_\alpha}^2 \equiv \sum_f \sigma_{\alpha(f)}^2.$$

A similar weighted sum can be used to optimally combine the estimators calculated for different time intervals \(^{[42]}\).

In the presence of correlated noise, the estimators are biased. The expected values are then

$$\langle \hat{\Omega}_\alpha(f) \rangle = \Omega_\alpha + \eta_\alpha(f),$$

where

$$\eta_\alpha(f) \equiv \frac{\Re [N_{12}(f)]}{\gamma(f)S_\alpha(f)}.$$

Here \(N_{12}(f) \equiv \frac{1}{2} \Re [\tilde{n}_1(f)\tilde{n}_2(f)]\) is the one-sided cross-spectral density (CSD) of the correlated noise contribution \(\tilde{n}_1, \tilde{n}_2\) to \(\tilde{s}_1, \tilde{s}_2\). The expression for the variance \(\sigma_{\hat{\Omega}_\alpha}^2(f)\) is unchanged in the presence of correlated noise provided \(|N_{12}(f)| \ll P_1(f), P_2(f)\). For the summed estimator \(\hat{\Omega}_\alpha\), we have

$$\langle \hat{\Omega}_\alpha \rangle = \Omega_\alpha + \eta_\alpha,$$

where

$$\eta_\alpha = \frac{\sum_f \sigma_{\hat{\Omega}_\alpha(f)}^2 \eta_\alpha(f)}{\sum_f \sigma_{\hat{\Omega}_\alpha(f)}^2}.$$
IV. METHODS FOR IDENTIFYING CORRELATED NOISE

A. Coherence calculation

Perhaps the simplest method for identifying correlated noise in the H1-H2 data is to calculate the magnitude squared coherence, \( \hat{\Gamma}_{12}(f) \equiv |\gamma_{12}(f)|^2 \), where

\[
\gamma_{12}(f) \equiv 2 \frac{\langle \hat{s}_1^*(f) \hat{s}_2(f) \rangle_N}{T \sqrt{\langle P_1(f) \rangle_N \langle P_2(f) \rangle_N}}. \tag{12}
\]

Here \( T \) denotes the duration of a single segment of data, and angle brackets \( \langle \rangle_N \) denotes an average over \( N \) segments used to estimate the CSD and PSDs that enter the expression for \( \gamma_{12} \). If there are no correlations (either due to noise or a GW signal) in the data, the expected value of \( \hat{\Gamma}_{12}(f) \) is equal to 1/\( N \). This method is especially useful at finding narrowband features that stick out above the expected 1/\( N \) level. Since we expect a SGWB to be broadband, with relatively little variation in the LIGO band (\( \sim 80–1000 \) Hz), most of these features can be attributed to instrumental and/or environmental correlations. We further investigate these lines with data from other PEM channels and once we confirm that they are indeed environmental/instrumental artifacts, we remove them from our analysis.

Plots of \( \hat{\Gamma}_{12}(f) \) for three different frequency resolutions are shown in Figs. 1 and 2 for two frequency bands, 80–160 Hz and 460–1000 Hz, respectively. In Fig. 1 note the relatively wide structure around 120 Hz, which is especially prominent in the bottom panel where the frequency resolution is 100 mHz. This structure arises from low-frequency noise (dominated by seismic and other mechanical noise) up-converting to frequencies around the 60 Hz harmonics via a bilinear coupling mechanism. While these coupling mechanisms are not fully understood, we reject the band from 102–126 Hz for our analysis, given the elevated correlated noise seen in this band. (A similar plot at slightly lower and higher frequencies shows similar noisy bands from 40–80 Hz and 160–200 Hz.) A closer look at the coherence also identifies smaller structures at 86–90 Hz, 100 Hz, 140–141 Hz, and 150 Hz. A follow-up analysis of PEM channels (which is discussed in more detail later) revealed that the grayed bands in Figs. 1 and 2 were highly contaminated with acoustic noise or by low-frequency seismic noise up-converting to frequencies around the 60 Hz harmonics via a bilinear coupling mechanism; so we rejected these frequency bands from subsequent analysis. As mentioned earlier, the 160–460 Hz band was not used in this analysis, because of similar acoustic and seismic contamination, as well as violin-mode resonances of the mirror-suspension wires (see Sec. [VI]).

As shown in Fig. 2, the coherence at high frequencies (460–1000 Hz) is relatively clean. The only evidence of narrowband correlated noise is in ±2 Hz bands around the 60 Hz power-line harmonics, and violin-mode resonances of mirror suspensions at 688.5 ± 2.8 Hz and 697 ± 3.1 Hz. The elevated coherence near 750 Hz at 100 mHz resolution is due to acoustic noise coupling to the GW channels. Notching the power-line harmonics and violin-mode resonances amounts to 100 mHz. This structure arises from low-frequency seismic noise up-converting to frequencies around the 60 Hz harmonics via a bilinear coupling mechanism. While these coupling mechanisms are not fully understood, we reject the band from 102–126 Hz for our analysis, given the elevated correlated noise seen in this band. (A similar plot at slightly lower and higher frequencies shows similar noisy bands from 40–80 Hz and 160–200 Hz.) A closer look at the coherence also identifies smaller structures at 86–90 Hz, 100 Hz, 140–141 Hz, and 150 Hz. A follow-up analysis of PEM channels (which is discussed in more detail later) revealed that the grayed bands in Figs. 1 and 2 were highly contaminated with acoustic noise or by low-frequency seismic noise up-converting to frequencies around the 60 Hz harmonics via a bilinear coupling mechanism; so we rejected these frequency bands from subsequent analysis. As mentioned earlier, the 160–460 Hz band was not used in this analysis, because of similar acoustic and seismic contamination, as well as violin-mode resonances of the mirror-suspension wires (see Sec. [VI]).
the removal of $\sim 9\%$ of the frequency bins over the entire high-frequency band.

### B. Time-shift analysis

A second method for identifying narrowband correlated noise is to time-shift the time-series output of one detector relative to that of the other detector before doing the cross-correlation analysis [43]. By introducing a shift of $\pm 1$ second, which is significantly larger than the correlation time for a broadband GW signal ($\sim 10$ ms, cf. Fig. 9), we eliminate broadband GW correlations while preserving narrowband noise features. Using segments of duration $T = 1$ s, we calculate the time-shifted estimators $\hat{\Omega}_{\alpha,TS}(f)$, variance $\sigma_{\hat{\Omega}_{\alpha,TS}(f)}^2$, and their ratio $\text{SNR}_{\hat{\Omega}_{\alpha,TS}(f)} \equiv \hat{\Omega}_{\alpha,TS}(f)/\sigma_{\hat{\Omega}_{\alpha,TS}(f)}$. The calibration and conditioning of the data is performed in exactly the same way as for the final search, which is described in detail in Secs. V and VI.

We excise any frequency bin with $|\text{SNR}_{\hat{\Omega}_{\alpha,TS}(f)}| > 2$ on the grounds that it is likely contaminated by correlated noise. This threshold was chosen on the basis of initial studies performed using playground data to understand the effectiveness of such cut. This criterion can be checked for different time-scales, such as weeks, months, or the entire data set. This allows us to identify transient effects on different time-scales, which may be diluted (and unobservable) when averaged over the entire data set.

### C. PEM coherence calculations

Another method for identifying correlated noise is to first try to identify the noise sources that couple into the individual detector outputs by calculating the coherence of $\hat{s}_1$ and $\hat{s}_2$ with various PEM channels $\hat{z}_I$:

$$\hat{\gamma}_{iI}(f) \equiv \frac{2}{T} \frac{\langle \hat{s}_1(f)\hat{z}_I(f) \rangle_N}{\sqrt{\langle P_i(f) \rangle_N \langle P_I(f) \rangle_N}}.$$  \hfill (13)

Here $i = 1, 2$ labels the detector outputs and $I$ labels the PEM channels. For our analysis we used 172 PEM channels located near the two detectors. In addition to the PEM channels, we used a couple of auxiliary channels associated with the stabilization of the frequency of the lasers used in the detectors, which potentially carry information about instrumental correlations between the two detectors. (Hereafter, the usage of the acronym PEM will also include these two auxiliary channels.)

The Fourier transforms are calculated for each minute of data ($T = 60$ s), and the average CSDs and PSDs are computed for extended time-periods—weeks, months, or the entire run. We then perform the following maximization over all PEM channels, for each frequency bin $f$, defining:

$$\hat{\gamma}_{12,PEM}(f) \equiv \max_I \Re [\hat{\gamma}_{1I}(f) \times \hat{\gamma}_{2I}^*(f)].$$  \hfill (14)

Note that by construction $\hat{\gamma}_{12,PEM}(f)$ is real.
As discussed in [44], \( g_{12,PEM}(f) \) is an estimate of the instrumental or environmental contribution to the coherence between the GW channels of H1 and H2. This estimate is only approximate, however, and potentially suffers from systematic errors for a few reasons. First, the PEM coverage of the observatory may be incomplete—i.e., there may be environmental or instrumental effects that are not captured by the existing array of PEMs. Second, some of the PEM channels may be correlated. Hence, a rigorous approach would require calculating a matrix of elements \( g_{ij}(f) \), and then inverting this matrix or solving a set of linear equations involving elements of \( g_{ij}(f) \). In practice, due to the large number of channels and the large amount of data, this is a formidable task. Instead, we simply maximize, frequency-by-frequency, over the contributions from different PEM channels and use this maximum as an estimate of the overall environmental contribution to \( g_{12}(f) \). Finally, these coherence methods do not take into account the nonlinear upconversion processes in which low-frequency disturbances, primarily seismic activity, excite higher-frequency modes in the instrument.

Since the measured signal-to-noise ratio for the estimator \( \hat{\Omega}_\alpha(f) \) can be written as

\[
\text{SNR}(f) = \sqrt{2T\Delta f} \Re [g_{12}(f)],
\]

we can simply approximate the contribution of the PEM channels to the stochastic GW signal-to-noise ratio as

\[
\text{SNR}_{PEM}(f) \equiv \sqrt{2T\Delta f} \hat{g}_{12,PEM}(f),
\]

remembering that \( \hat{g}_{12,PEM}(f) \) is real. The PEM contribution to the estimators \( \hat{\Omega}_\alpha(f) \) is then

\[
\hat{\Omega}_{\alpha,PEM}(f) \equiv \text{SNR}_{PEM}(f) \sigma_{\hat{\Omega}_\alpha}(f)
\]

where \( \sigma_{\hat{\Omega}_\alpha}(f) \) is the statistical uncertainty defined by Eq. 6.

We can use the PEM coherence calculations in two complementary ways. First, we can identify frequency bins with particularly large instrumental or environmental contributions by placing a threshold on \( |\text{SNR}_{PEM}(f)| \) and exclude them from the analysis. Second, the frequency bins that pass this data-quality cut may still contain some residual environmental contamination. We can estimate at least part of this residual contamination by using \( \hat{\Omega}_{\alpha,PEM}(f) \) for the remaining frequency bins.

As part of the analysis procedure, we were able to identify the PEM channels that were responsible for the largest coherent noise between the GW channels in H1 and H2 for each frequency bin. For both the low and high frequency analyses, microphones and accelerometers in the central building near the beam splitters of each interferometer registered the most significant noise. Within approximately 1 Hz of the 60-Hz harmonics, magnetometers and voltage line monitors registered the largest correlated noise, but these frequencies were already removed from the analysis due to the significant coherence (noise) level at these frequencies, as mentioned in Sec. [IV A]

![Figure 3](image.png)

**FIG. 3.** (Color online) Comparison of the (absolute value of the) SNRs calculated by the PEM-coherence and the time-shift techniques. The vertical dotted lines indicate the frequency bands used for the low (80–160 Hz; black dotted lines) and high (460–1000 Hz; magenta dotted lines) frequency analyses. Note that \( \text{SNR}_{\alpha,TS}(f) \) is a true signal-to-noise ratio, so values \( \geq 2 \) are dominated by random statistical fluctuations. \( \text{SNR}_{\alpha,PEM}(f) \), on the other hand, is an estimate of the PEM contribution to the signal-to-noise ratio, so values even much lower than 2 are meaningful measurements (i.e., they are not statistical fluctuations). The two methods agree very well in identifying contaminated frequency bins or bands. Note that both methods indicate that the 80–160 Hz and 460–1000 Hz bands have relatively low levels of contamination.

**D. Comparing PEM-coherence and time-shift methods**

Figure 3 shows a comparison of the SNRs calculated by the PEM-coherence and time-shift methods. The agreement between these two very different techniques in identifying contaminated frequency bins (those with \( |\text{SNR}| \geq \alpha \) few) is remarkably good, which is an indication of their robustness and effectiveness. Moreover, Fig. 3 shows that the frequency region between 200 Hz and 460 Hz is particularly contaminated by environmental and/or instrumental effects. Hence, in this analysis we focus on the low-frequency region (80–160 Hz) which is the most sensitive to cosmological backgrounds (i.e., spectral index \( \alpha = 0 \)), and on the high-frequency region (460–1000 Hz) which is less contaminated and more suitable for searches for astrophysically-generated backgrounds (e.g., \( \alpha = 3 \)).

We emphasize that the PEM channels only monitor the instrument and the environment, and are not sensitive to GWs. Similarly, the time-shift analysis, with a time-shift of \( \pm 1 \) second, is insensitive to broad-band GW signals. Hence, any data-quality cuts based on the PEM and time-shift studies will not affect the astrophysical signatures in the data—i.e., they do not bias our estimates of the amplitude of a SGWB.
E. Other potential non-astrophysical sources of correlation

We note that any correlations that are produced by environmental signals that are not detected by the PEM sensors will not be detected by the PEM-coherence technique. Furthermore, if such correlations, or correlations from a non-environmental source, are broadband and flat (i.e., do not vary with frequency over our band), they will not be detected by either the PEM-coherence or the time-shift method. One potential source of broadband correlation between the two GW channels is the data acquisition system itself. We investigated this possibility by looking for correlations between 153 channel pairs that had no physical reason to be correlated. We found no broadband correlations, although we did find an unexplained narrow-band correlation at 281.5 Hz between 10 of 153 channel pairs. Note that 281.5 Hz is outside of the frequency bands analyzed in this study.

We also examined the possibility of correlations between the H1 and H2 detectors being generated by scattered light. We considered two mechanisms: first, light scattered from one detector affecting the other detector, and second, light from both detectors scattering off of the same site and returning to the originating detectors. We did not observe, and do not expect to observe, the first mechanism because the frequencies of the two lasers, while very stable, may differ by gigahertz. If light from one interferometer scatters into the main beam of the other, it will likely be at a very different frequency and will not produce signals in our 8 kHz band when it beats against the reference light for that interferometer.

Nevertheless, we checked for a correlation produced by light from one detector entering the other by looking for the calibration signals [5] injected into one detector in the signal of the second detector. During S5, the following calibration line frequencies were injected into H1 and H2: 46.70 Hz, 393.10 Hz, 1144.30 Hz (H1) and 54.10 Hz, 407.30 Hz, 1159.7 Hz (H2). We note here that all those frequencies are 1% of the total frequency notches; hence it is safe to assume that we had sufficient PEM coverage throughout our analysis period.

We also investigated this possibility by looking for correlations between 153 channel pairs that had no physical reason to be correlated. We found no broadband correlations, although we did find an unexplained narrow-band correlation at 281.5 Hz between 10 of 153 channel pairs. Note that 281.5 Hz is outside of the frequency bands analyzed in this study.

We addressed the potential of correlations from unmonitored environmental signals by searching for coupling sites four times over the course of the run by injecting large but localized acoustic, seismic, magnetic and RF signals. New sensors were installed at the two coupling sites that had the least coverage. However, we found that the new sensors, even after scaling up to the full analysis period, contribute less than 1% of the total frequency notches; hence it is safe to assume that we had sufficient PEM coverage throughout our analysis period.

We also investigated the possibility of correlations between the H1 and H2 detectors being generated by scattered light. We considered two mechanisms: first, light scattered from one detector affecting the other detector, and second, light from both detectors scattering off of the same site and returning to the originating detectors. We did not observe, and do not expect to observe, the first mechanism because the frequencies of the two lasers, while very stable, may differ by gigahertz. If light from one interferometer scatters into the main beam of the other, it will likely be at a very different frequency and will not produce signals in our 8 kHz band when it beats against the reference light for that interferometer.

Nevertheless, we checked for a correlation produced by light from one detector entering the other by looking for the calibration signals [5] injected into one detector in the signal of the second detector. During S5, the following calibration line frequencies were injected into H1 and H2: 46.70 Hz, 393.10 Hz, 1144.30 Hz (H1) and 54.10 Hz, 407.30 Hz, 1159.7 Hz (H2). We note here that all those frequencies are 1% of the total frequency notches; hence it is safe to assume that we had sufficient PEM coverage throughout our analysis period.

We also investigated the possibility of correlations between the H1 and H2 detectors being generated by scattered light. We considered two mechanisms: first, light scattered from one detector affecting the other detector, and second, light from both detectors scattering off of the same site and returning to the originating detectors. We did not observe, and do not expect to observe, the first mechanism because the frequencies of the two lasers, while very stable, may differ by gigahertz. If light from one interferometer scatters into the main beam of the other, it will likely be at a very different frequency and will not produce signals in our 8 kHz band when it beats against the reference light for that interferometer.

Nevertheless, we checked for a correlation produced by light from one detector entering the other by looking for the calibration signals [5] injected into one detector in the signal of the second detector. During S5, the following calibration line frequencies were injected into H1 and H2: 46.70 Hz, 393.10 Hz, 1144.30 Hz (H1) and 54.10 Hz, 407.30 Hz, 1159.7 Hz (H2). We note here that all those frequencies are 1% of the total frequency notches; hence it is safe to assume that we had sufficient PEM coverage throughout our analysis period.

We also investigated the possibility of correlations between the H1 and H2 detectors being generated by scattered light. We considered two mechanisms: first, light scattered from one detector affecting the other detector, and second, light from both detectors scattering off of the same site and returning to the originating detectors. We did not observe, and do not expect to observe, the first mechanism because the frequencies of the two lasers, while very stable, may differ by gigahertz. If light from one interferometer scatters into the main beam of the other, it will likely be at a very different frequency and will not produce signals in our 8 kHz band when it beats against the reference light for that interferometer.

We also investigated the possibility of correlations between the H1 and H2 detectors being generated by scattered light. We considered two mechanisms: first, light scattered from one detector affecting the other detector, and second, light from both detectors scattering off of the same site and returning to the originating detectors. We did not observe, and do not expect to observe, the first mechanism because the frequencies of the two lasers, while very stable, may differ by gigahertz. If light from one interferometer scatters into the main beam of the other, it will likely be at a very different frequency and will not produce signals in our 8 kHz band when it beats against the reference light for that interferometer.
STEP 3: We perform a search for transient excess power in the data using the wavelet-based Kleine Welle algorithm [45], which was originally designed for detecting GW bursts. This algorithm is applied to the output of both detectors, producing a list of triggers for each detector. We then search the two trigger lists and reject any segment that contains transients with Kleine Welle significance larger than 50 in either of the two detectors. The value of 50 is a conservative threshold, chosen based on other studies done on the distribution of such triggers in S5 [32].

STEP 4: Having determined the reasonably good frequency bands, we then calculate $\hat{\Omega}_\alpha$ and its uncertainty $\sigma_{\hat{\Omega}_\alpha}$ summed over the whole band, cf. Eq. 7. The purpose of this calculation is to perform another level of data-quality selection in the time-domain by identifying noisy segments of 60 s duration. It is similar to the non-stationarity cut used in the previous analyses [24, 28, 46] where we remove time segments whose $\sigma_{\hat{\Omega}_\alpha}$ differs by a pre-determined amount, from that calculated by averaging over two neighboring segments. Here we use a 20% threshold on the difference. The combination of the time-domain data quality cuts described in Steps 1–4 removed about 22% of the available S5 H1-H2 data.

STEP 5: After identifying and rejecting noisy time segments and frequency bins using Steps 1–4, we then use the time-shift and the PEM-coherence methods described in Secs. IV B and IV C to identify any remaining contaminated frequency bins. To remove bad frequency bins, we split the S5 dataset into week-long periods and for each week, we reject any frequency bin for which either $|\text{SNR}_{\Omega_{\alpha,\text{TS}}}(f)|$ or $|\text{SNR}_{\Omega_{\alpha,\text{PEM}}}(f)|$ exceeds a pre-determined threshold in the given week, the corresponding month, or in the entire S5 dataset. This procedure generates (different) sets of frequency notchings for each week of the S5 dataset. In the analysis we use two different sets of SNR threshold values for the cut, which are further described in Sec. VI.

Figure 4 is a spectrogram of $\text{SNR}_{\Omega_{\alpha,\text{PEM}}}$ for the 80–160 Hz band for all weeks in S5; the visible structure represents correlated noise between H1 and H2, which was identified and subsequently excluded from the analysis by the H1-H2 coherence, time-shift, and PEM-coherence measurements.

Note that previous stochastic analyses using LIGO data [24, 28, 46] followed only steps 1, 2 and 4. Steps 3 and 5 were developed for this particular analysis.

Having defined the time-segments and frequency-bins to be rejected in each week of the S5 data, we proceed with the calculation of the estimators and standard errors, $\Omega_{\alpha}(f)$ and $\sigma_{\hat{\Omega}_\alpha}(f)$, in much the same manner as in previous searches for isotropic stochastic backgrounds [24, 28, 47]. The data is divided into $T = 60 \text{ s}$ segments, decimated to 1024 Hz for the low-frequency analysis and 4096 Hz for the high-frequency analysis, and high-pass filtered with a 6th order Butterworth filter with 32 Hz knee frequency. Each analysis segment is Hann-windowed, and to recover the loss of signal-to-noise ratio due to Hann-windowing, segments are 50% overlapped. Estimators and standard errors for each segment are evaluated with a $\Delta f = 0.25 \text{ Hz}$ frequency resolution, using the frequency mask of the week to which the segment belongs. A weighted average is performed over all segments and all frequency bins, with inverse variances, as in Eq. 7 but properly accounting for overlapping.

VI. ANALYSIS RESULTS

The analysis is separated into two parts corresponding to searches for SGWBs with spectral index $\alpha = 0$ and $\alpha = 3$ as described in Sec. III. Since the strain output of an interferometer due to GWs is $S_{gw}(f) \propto f^{\alpha-3}$ (see Eq. 3), the case $\alpha = 0$ is dominated by low frequencies while $\alpha = 3$ is independent of frequency. Since for $\alpha = 3$ there is no preferred frequency band, and since previous analyses [46] for stochastic backgrounds with $\alpha = 3$ considered only high frequencies, we also used only high frequencies for the $\alpha = 3$ case. Thus, the two cases of $\alpha = 0$ and $\alpha = 3$ correspond to the analysis of the low and high-frequency bands, respectively. In this section, we present the results of the analyses in the two different frequency bands as defined in Sec. IV D corresponding to the two different values of $\alpha$.

To illustrate the effect of the various noise removal methods described in the previous two sections, we give the results as different stages of cuts are applied to the data (see Table I). The threshold value used at stage III comes from an initial study performed using playground data to understand the effectiveness of the PEM-coherence method in finding problematic frequency bins in the H1-H2 analysis, and hence those results are considered as blind analysis results. But a post-unblinding study showed that we could lower the $\text{SNR}_{\text{PEM}}$ threshold to values as low as 0.5 (for low-frequency) and 1 (for high-frequency), which are used at stage IV. These post-unblinding results are used in the final upper-limit calculations. For threshold values $< 0.5$ (low-frequency) or $< 1$ (high-frequency), the PEM-coherence contribution, $\hat{\Omega}_{\alpha,\text{PEM}}$, varies randomly as the threshold is changed indicating the statistical noise limit of the PEM-coherence method.

A. High-frequency results

We performed the high-frequency analysis with spectral index $\alpha = 3$, and reference frequency $f_{\text{ref}} = 900$ Hz. Tables II and III summarize the results after applying several stages of noise removal as defined in Table I. Table II applies to the full analysis band, 460–1000 Hz; Table III gives the results for 5 separate sub-bands. The values of the estimator, $\hat{\Omega}_3$, the PEM-coherence contribution to the estimator, $\hat{\Omega}_{3,\text{PEM}}$, and the statistical uncertainty, $\sigma_{\hat{\Omega}_3}$, are given for each band and each stage of noise removal. Also given is the ratio of the standard deviation of the values of the inverse Fourier transform of $\hat{\Omega}_3(f)$ to the statistical uncertainty $\sigma_{\hat{\Omega}_3}$, which is a measure of excess residual correlated noise. In the absence of correlated noise, we expect the distribution of data points in
the inverse Fourier transform of $\hat{\Omega}_3(f)$ to follow a Gaussian distribution with mean 0 and std $\sigma_{\hat{\Omega}_3}$. Hence a ratio $\gg 1$ is a sign of excess cross-correlated noise. The left column gives the ratio of the standard deviation of the values of the inverse Fourier transform of $\hat{\Omega}_3(f)$ to the statistical uncertainty $\sigma_{\hat{\Omega}_3}$. As described in Sec. VI A a ratio much $\gg 1$ is a sign of excess cross-correlated noise. The PEM-coherence estimate on stage I also excludes frequencies (including 60 Hz harmonics) and time segments similar to stages II-IV.

Figure 4. (Color online) Spectrograms displaying the absolute value of $\text{SNR}_{\hat{\Omega}_3,\text{PEM}}(f)$ for 80–160 Hz (left) and 460–1000 Hz (right) as a function of the week in S5. The horizontal dark (blue) lines correspond to initial frequency notches as described in STEP 2 (Sec. V) and vertical dark (blue) lines correspond to unavailability of data due to detector downtime. The large SNR structures seen in the plots were removed from the low- and high-frequency analyses.

The horizontal dark (blue) bands correspond to initial frequency notches as described in STEP 2 (Sec. V) and vertical dark (blue) lines correspond to unavailability of data due to detector downtime. The large SNR structures seen in the plots were removed from the low- and high-frequency analyses.

The middle dark (blue) line in the plot of the inverse Fourier transform of $\hat{\Omega}_3(f)$ indicates the presence of excess correlated noise, which shows up as visible structure in the plot of the inverse Fourier transform of $\hat{\Omega}_3(f)$ (for example, see the right hand plots in Fig. 5). We see that this ratio decreases for the full 460–1000 Hz band and for each sub-band with every stage of data cleanup indicating the effectiveness of PEM-coherence SNR cut. We also note that the values listed in Tables II-IV are the zero lag values of $\hat{\Omega}_3$ in the corresponding inverse Fourier transform plots.

Figure 5 is devoted entirely to the noisiest sub-band, 628–733 Hz. The left column of plots shows $\hat{\Omega}_3(f)$ and $\hat{\Omega}_3,\text{PEM}(f)$, with black lines denoting the statistical error bar $\pm \sigma_{\hat{\Omega}_3}$. Here we can clearly see the effectiveness of noise removal through the four stages discussed above. Note the lack of structure near zero-lag in the final inverse Fourier transform of $\hat{\Omega}_3(f)$ which is consistent with no correlated noise. Figure 6 is a similar plot for the full 460–1000 Hz band, showing the results after the final stage of cuts. Again note the lack of significant structure near zero-lag in the inverse Fourier transform of $\hat{\Omega}_3(f)$. Figure 7 (left panel) shows how the final estimate, $\hat{\Omega}_3$, summed over the whole band, evolves over the course of the run after the final stage of cuts. The smoothness of that plot (absence of any sharp rise or fall after the accumulation of sufficient data i.e., one month) indicates that no particular time period dominates our

<table>
<thead>
<tr>
<th>Stage</th>
<th>High-frequency analysis</th>
<th>Low-frequency analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Steps</td>
<td>% of data vetoed</td>
</tr>
<tr>
<td>I</td>
<td>Step 1</td>
<td>8.51</td>
</tr>
<tr>
<td>II</td>
<td>Steps 1–4</td>
<td>35.88</td>
</tr>
<tr>
<td>III</td>
<td>Steps 1–5 with $</td>
<td>\text{SNR}_{\text{PEM}}</td>
</tr>
<tr>
<td>IV</td>
<td>Steps 1–5, with $</td>
<td>\text{SNR}_{\text{PEM}}</td>
</tr>
</tbody>
</table>

TABLE I. Definition of various stages of noise removal for the high and low-frequency analyses in terms of the analysis steps described in Sec. V. Here stage III corresponds to the blind analysis and stage IV to the post-unblinding analysis. The percentage of data vetoed accounts for both the time segments and frequency bins excluded from the analysis. In calculating veto percentage, the analyses with non-colocated LIGO detectors only accounts for the time segments excluded from the analyses and is the reason for the large numbers we see in the last column compared to other LIGO analyses.

<table>
<thead>
<tr>
<th>Stage</th>
<th>$\hat{\Omega}_3$</th>
<th>$\hat{\Omega}_{3,\text{PEM}}$</th>
<th>$\sigma_{\hat{\Omega}_3}$</th>
<th>$\text{std}/\sigma_{\hat{\Omega}_3}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>77.5</td>
<td>$-3.05^\dagger$</td>
<td>2.82</td>
<td>20.5</td>
</tr>
<tr>
<td>II</td>
<td>$-2.17$</td>
<td>$-3.62$</td>
<td>3.24</td>
<td>1.18</td>
</tr>
<tr>
<td>III</td>
<td>$-4.11$</td>
<td>$-4.30$</td>
<td>3.59</td>
<td>1.04</td>
</tr>
<tr>
<td>IV</td>
<td>$-1.29$</td>
<td>$-2.38$</td>
<td>3.64</td>
<td>1.01</td>
</tr>
</tbody>
</table>

TABLE II. Results for the H1-H2 high-frequency analysis (460–1000 Hz) after various stages of noise removal were applied to the data. The estimates $\hat{\Omega}_3$, PEM-coherence contribution, $\hat{\Omega}_{3,\text{PEM}}$ and $\sigma_{\hat{\Omega}_3}$ are calculated assuming $H_0 = 68 \text{ km/s/Mpc}$. $\sigma_{\hat{\Omega}_3}$ is the statistical uncertainty in $\hat{\Omega}_3$. The last column gives the ratio of the standard deviation of the values of the inverse Fourier transform of $\hat{\Omega}_3(f)$ to the statistical uncertainty $\sigma_{\hat{\Omega}_3}$. As described in Sec. VI A a ratio much $\gg 1$ is a sign of excess cross-correlated noise. The PEM-coherence estimate on stage I also excludes frequencies (including 60 Hz harmonics) and time segments similar to stages II-IV.
TABLE IV. Similar to Table II but for the low-frequency analysis (80–160 Hz with spectral index \(\alpha = 0\)). The different rows give the results after various stages of noise removal were applied to the data.

<table>
<thead>
<tr>
<th>Band (Hz)</th>
<th>Stage</th>
<th>(\hat{\Omega}_3) ((x \times 10^{-6}))</th>
<th>(\hat{\Omega}_{3,\text{PEM}}) ((x \times 10^{-6}))</th>
<th>(\sigma_{\hat{\Omega}_3}) ((x \times 10^{-6}))</th>
<th>(\text{std/}\sigma_{\hat{\Omega}_3})</th>
</tr>
</thead>
<tbody>
<tr>
<td>460–537</td>
<td>I</td>
<td>−7.28</td>
<td>−0.22</td>
<td>4.48</td>
<td>5.40</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>−2.17</td>
<td>−0.24</td>
<td>5.08</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>−0.60</td>
<td>−1.23</td>
<td>5.68</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>IV</td>
<td>−0.34</td>
<td>−1.23</td>
<td>5.09</td>
<td>0.97</td>
</tr>
<tr>
<td>537–628</td>
<td>I</td>
<td>163</td>
<td>−2.28</td>
<td>5.46</td>
<td>24.0</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>14.7</td>
<td>−2.46</td>
<td>6.32</td>
<td>1.08</td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>8.83</td>
<td>−2.00</td>
<td>6.96</td>
<td>1.02</td>
</tr>
<tr>
<td></td>
<td>IV</td>
<td>8.56</td>
<td>−1.98</td>
<td>7.03</td>
<td>1.02</td>
</tr>
<tr>
<td>628–733</td>
<td>I</td>
<td>512</td>
<td>−16.7</td>
<td>7.33</td>
<td>35.9</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>−33.2</td>
<td>−20.5</td>
<td>8.52</td>
<td>1.37</td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>−37.0</td>
<td>−16.3</td>
<td>9.20</td>
<td>1.21</td>
</tr>
<tr>
<td></td>
<td>IV</td>
<td>−26.5</td>
<td>−5.88</td>
<td>9.66</td>
<td>1.12</td>
</tr>
<tr>
<td>733–856</td>
<td>I</td>
<td>−397</td>
<td>−1.77</td>
<td>8.32</td>
<td>23.0</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>−4.44</td>
<td>−2.24</td>
<td>9.49</td>
<td>1.67</td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>−5.29</td>
<td>−6.40</td>
<td>11.0</td>
<td>1.04</td>
</tr>
<tr>
<td></td>
<td>IV</td>
<td>2.76</td>
<td>−3.91</td>
<td>11.3</td>
<td>0.98</td>
</tr>
<tr>
<td>856–1000</td>
<td>I</td>
<td>89.2</td>
<td>4.63</td>
<td>10.6</td>
<td>3.37</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>2.44</td>
<td>4.63</td>
<td>12.0</td>
<td>1.02</td>
</tr>
<tr>
<td></td>
<td>III</td>
<td>0.004</td>
<td>−1.47</td>
<td>13.2</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td>IV</td>
<td>0.21</td>
<td>−1.41</td>
<td>13.2</td>
<td>1.01</td>
</tr>
</tbody>
</table>

TABLE III. Same as Table II but for 5 separate sub-bands of 460–1000 Hz.

<table>
<thead>
<tr>
<th>Stage</th>
<th>(\hat{\Omega}_3) ((x \times 10^{-6}))</th>
<th>(\hat{\Omega}_{3,\text{PEM}}) ((x \times 10^{-6}))</th>
<th>(\sigma_{\hat{\Omega}_3}) ((x \times 10^{-6}))</th>
<th>(\text{std/}\sigma_{\hat{\Omega}_3})</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>6.17</td>
<td>−0.39</td>
<td>0.44</td>
<td>5.90</td>
</tr>
<tr>
<td>II</td>
<td>−1.71</td>
<td>−0.78</td>
<td>0.63</td>
<td>1.80</td>
</tr>
<tr>
<td>III</td>
<td>−1.57</td>
<td>−0.84</td>
<td>0.79</td>
<td>1.64</td>
</tr>
<tr>
<td>IV</td>
<td>−0.26</td>
<td>−0.29</td>
<td>0.85</td>
<td>1.63</td>
</tr>
</tbody>
</table>

B. Low frequency results

We now repeat the analysis of the previous subsection but for the low-frequency band, 80–160 Hz with spectral index \(\alpha = 0\) and \(f_{\text{ref}} = 100\) Hz. Table IV summarizes the results for the low-frequency analysis after applying several stages of noise removal as defined in Table II. Figure 8 shows the results obtained by applying the noise removal cuts in four stages. The left column of plots contain the estimators, \(\hat{\Omega}_3(f)\) and \(\hat{\Omega}_{3,\text{PEM}}(f)\), with lines denoting the statistical error bar \(\pm \sigma_{\hat{\Omega}_3}(f)\).

In contrast to the high-frequency analysis (compare Figs. 6 and 8) there is still much structure in the inverse Fourier transform of \(\hat{\Omega}_3(f)\) around zero-lag even after the final stage of noise removal cuts were applied. In addition, the PEM-coherence contribution to the estimator, \(\hat{\Omega}_{3,\text{PEM}}(f)\), displays much of the structure observed in \(\hat{\Omega}_3(f)\). Both of these observations suggest contamination from residual correlated instrumental or environmental noise that was not excluded by the noise removal methods. Figure 7 (right panel) shows how the final estimate, \(\hat{\Omega}_3\), evolves over the course of the run after the final stage of cuts. We note here that even though \(\hat{\Omega}_3\) (last entry in Table IV) is consistent with zero (within 2\(\sigma\)), its estimate at other non-zero lags vary strongly as shown in Fig. 8 (lower right). This indicates the presence of residual correlated noise after all the time-shift and PEM-coherence noise removal cuts are applied.

C. Hardware and software injections

We validated our analysis procedure by injecting simulated stochastic GW signals into the strain data of the two detectors. Both hardware and software injections were performed. Hardware injections are performed by physically moving the interferometer mirrors coherently between interferometers. In this case the artificial signals were limited to short durations and relatively large amplitudes. The data from these hardware injection times were excluded from the analyses described above, as noted in Sec. V, Step 1. Software injections are conducted by adding a simulated GW signal to the interferometer data, in which case they could be long in duration and relatively weak in amplitude. During S5 there was one stochastic signal hardware injection when both H1 and H2 were operating in coincidence. A stochastic background signal with spectral index \(\alpha = 0\) and amplitude \(\hat{\Omega}_3 = 6.56 \times 10^{-3}\) was injected for approximately 3 hours. In performing the analysis, frequency bins were excluded based on the standard H1-H2 coherence calculations. No additional frequency bins were removed using SNRPEM. The recovered signal was \(\hat{\Omega}_3 = (7.39 \pm 1.1) \times 10^{-3}\), which is consistent with the injected amplitude. Due to the spectral index used for the injection (\(\alpha = 0\)), the recovery analysis was performed using only the low frequency band. We also performed a software injection in the high frequency band with an amplitude \(\hat{\Omega}_3 = 5.6 \times 10^{-3}\), and we recovered it successfully. Figure 9 shows the spectrum of the recovered \(\hat{\Omega}_3(f)\) and its inverse Fourier transform.

VII. ASSESSING THE RESIDUAL CORRELATED NOISE

After applying the full noise removal procedure, the high-frequency band appears clean whereas the low-frequency band exhibits evidence of residual correlated noise. In order to interpret the implications of these two very different results, we introduce a general procedure for determining whether a stochastic measurement is sufficiently well-understood to...
FIG. 5. Plots of $\hat{\Omega}_3(f)$ and $\hat{\Omega}_{3,\text{PEM}}(f)$ (left), and the inverse Fourier transform of $\hat{\Omega}_3(f)$ (right), for the (noisiest) 628–733 Hz sub-band after various stages of noise removal were applied to the data. The four rows correspond to the four different stages of cleaning defined in Table I. (The top right plot has y-axis limits 13 x greater than the other three.)
yield an astrophysical interpretation. While our immediate concern is to provide a framework for interpreting the two results presented here, we aim to give a comprehensive procedure that can be applied generally, to both co-located and non-co-located detectors. In this spirit, this section is organized as follows: first, we present a general framework for interpreting stochastic measurements; then we discuss how it can be applied to (familiar) results from non-co-located detectors; and finally we apply the framework to our present results.

To determine whether a result can be interpreted as a constraint on the SGWB, we consider the following three criteria:

1. We have accounted for all known noise sources through either direct subtraction, vetoing, and/or proper estimation of systematic errors.

2. Having accounted for known noise sources, we do not observe evidence of residual noise that is inconsistent with our signal and noise models.

3. To the best of our knowledge, there is no plausible mechanism by which broadband correlated noise might be lurking beneath the uncorrelated noise at a level comparable to the GW signal we are trying to measure.

If an analysis result does not meet these criteria, then we conservatively place a bound on the sum of the GW signal and the residual correlated noise. If a result meets all the criteria, then we present astrophysical bounds on just the GW signal.

Let us now examine these criteria in the context of previous results using the non-co-located LHO and LLO detectors [28]. Criterion #1 was satisfied by identifying and removing instrumental lines attributable to known instrumental artifacts such as power lines and violin resonances. Criterion #2 was satisfied by creating diagnostic plots, e.g., showing $\hat{\Omega}_3$ vs. lag (the delay time between the detectors; see Fig. 5), which demonstrated that the measurement was consistent with uncorrelated noise (and no GW signal). Criterion #3 was satisfied by performing order-of-magnitude calculations for plausible sources.
FIG. 8. Plots of $\hat{\Omega}_0(f)$ and $\hat{\Omega}_{0,\text{PEM}}(f)$ (left), and the inverse Fourier transform of $\hat{\Omega}_0(f)$ (right) for the 80–160 Hz band after various stages of noise removal were applied to the data. The four rows correspond to the four different stages of cleaning defined in Table I.
of correlated noise for LHO-LLO including electromagnetic phenomena, and finding that they were too small to create broadband correlated noise at a level that is important for initial LIGO.

Next, we consider how the criteria might be applied to future measurements with non-co-located detectors. During the advanced detector era, correlated noise from Schumann resonances may constitute a source of correlated noise at low frequencies $\lesssim 200 \text{ Hz}$, even for widely separated detector pairs such as LHO-LLO \cite{29,30}. While it may be possible to mitigate this potential correlated noise source through commissioning of the detectors to minimize magnetic coupling, or failing that, through a noise subtraction scheme, we consider the possibility that residual correlated noise is observed. In this scenario, we could still aim to satisfy criteria #2 and #3 by using magnetometer measurements to construct a correlated noise budget, which could then be used to interpret the results.

Finally, we consider how the criteria apply to the measurements presented in this paper. The high-frequency analysis meets criteria 1 and 2 as we did not observe residual noise inconsistent with our noise models (see Fig. 6). We did observe residual noise for the low frequency analysis (see Fig. 8), but it was consistent with a preliminary noise model, based on measured acoustic coupling and microphone signals (most of the channels identified by the PEM coherence method were either microphones or accelerometers placed on optical tables that were susceptible to acoustic couplings). While the bands that were acoustically loudest (containing certain electronics fans) were vetoed, the acoustic coupling in between the vetoed bands was high enough to produce a residual signal. We did not further develop the noise model to meet criterion 1 because, with the systematic error from acoustic coupling, the astrophysical limit would not have improved on values we have reported previously \cite{28,48}. For this reason, we do not present an astrophysical limit for the low frequency band.

We addressed criterion #3 in two ways. First, by investigating mechanisms that might produce un-monitored broad-band correlations between detectors, such as the study of correlations introduced by the shared data acquisition system, the study of correlations introduced by light scattering, and PEM coverage studies described in Sec. IV E.

The second type of coherence feature was associated with bilinear coupling of low frequency ($< 15 \text{ Hz}$) seismic motion and harmonics of 60 Hz, producing side-band features around the harmonics that were similar to the features in the 0–15 Hz seismic band. Coherence of side-band features was expected since the coherence length of low-frequency seismic signals was greater than the distance separating sensitive parts of the two interferometers at the vertex station, and the seismic isolation of the interferometers was minimal below 10 Hz.

In conclusion, we found no peaks or features in the coherence spectrum for the two GW channels that were inconsistent with linear acoustic coupling or bilinear coupling of low frequency seismic noise and 60 Hz harmonics at the vertex station. Neither of these mechanisms is capable of producing broad-band coherence that is not well monitored by the PEM system. Therefore, for the high frequency analysis, we satisfy the three criteria for presenting astrophysical bounds on just the GW signal.

We also identified the sources of most of the features between 80 and 400 Hz. For many of the spectral peaks, in addition to coherence between the GW channels, there was also coherence between the individual GW channels and the accelerometer and microphone signals from the vertex area shared by both detectors. The coupling was consistent with the measured coupling of acoustic signals to the detectors. Most of these features were traced to electronics cooling fans in specific power supply racks in the vertex station by comparing coherence spectra to spectra for accelerometers mounted temporarily on each of the electronics racks. The features were produced at harmonics of the fan rotation frequencies.
A. Upper-limits

Since there is no evidence of significant residual noise contaminating the high-frequency data after applying the full set of cuts, we set a 95% confidence-level Bayesian upper-limit on $\Omega_3$. We use the previous high-frequency upper limit $\Omega_3 < 0.35$ (adjusted for $H_0 = 0.68$ km/s/Mpc) from the LIGO S5 and Virgo VSR1 analysis [46] as a prior and assume a flat distribution for $\Omega_3$ from 0 to 0.35. We also marginalize over the calibration uncertainty for the individual detectors (10.2% and 10.3% for H1 and H2, respectively). In order to include in our calculation the PEM estimate of residual contamination, we take $\sigma^2 + \Omega^2_{3,PEM}$ as our total variance. We note here that the estimated $\Omega_{3,PEM}$ is within the observed $\sigma_{\Omega_3}$, i.e., we observe no evidence of excess environmental contamination and the above quadrature addition increases the limit by $\sim 20\%$. The final result is $\Omega_3 < 7.7 \times 10^{-4}$ for the frequency band 460-1000 Hz, which is an improvement by a factor of $\sim 180$ over the recent S6/VSR2-3 result [48]. All of the above $\sim 180$ factor improvement comes from the nearly-unity overlap reduction function of the co-located Hanford detectors. In fact, all other data being same, if we were to consider the H2 detector to not be located at Hanford but instead at the LIGO Livingston site yields an upper limit that is worse by a factor of $\sim 1.7$ than the S6/VSR2-3 result. Most of this difference of $\sim 1.7$ comes from the improved sensitivities of S6/VSR2-3 detectors compared to S5 H1-H2 detectors. Upper-limits for the five separate sub-bands of the high-frequency analysis are given in Table V.

<table>
<thead>
<tr>
<th>Band (Hz)</th>
<th>95% CL UL ($\times 10^{-3}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>460–1000</td>
<td>0.77</td>
</tr>
<tr>
<td>460–537</td>
<td>1.11</td>
</tr>
<tr>
<td>537–628</td>
<td>2.12</td>
</tr>
<tr>
<td>628–733</td>
<td>1.18</td>
</tr>
<tr>
<td>733–856</td>
<td>2.53</td>
</tr>
<tr>
<td>856–1000</td>
<td>2.61</td>
</tr>
</tbody>
</table>

TABLE V. 95% confidence level upper-limits for the the full band (460–1000 Hz) and for five separate sub-bands.

As mentioned in Sec. IVB, the structure in the inverse Fourier transform plots of Fig. V suggests contamination from residual correlated noise for the low-frequency analysis and hence we do not set any upper-limit on $\Omega_0$ using the low-frequency band 80-160 Hz.

VIII. SUMMARY AND PLANS FOR FUTURE ANALYSES

In this paper, we described an analysis for a SGWB using data taken by the two co-located LIGO Hanford detectors, H1 and H2, during LIGO’s fifth science run. Since these detectors share the same local environment, it was necessary to account for the presence of correlated instrumental and environmental noise. We applied several noise identification and mitigation techniques to reduce contamination and to estimate the bias due to any residual correlated noise. The methods proved to be useful in cleaning the high-frequency band. But not enough in the low-frequency band.

In the 80 – 160 Hz band, we were unable to sufficiently mitigate the effects of correlated noise, and hence we did not set any limit on the GW energy density for $\alpha = 0$. For the 460 – 1000 Hz band, we were able to mitigate the effects of correlated noise, and so we placed a 95% C.L. upper limit on the GW energy density alone in this band of $\Omega_3 < 7.7 \times 10^{-4}$. This limit improves on the previous best limit in the high-frequency band by a factor of $\sim 180$ [48]. Figure 10 shows upper limits from current/past SGWB analyses, as well as limits from various SGWB models, and projected limits using Advanced LIGO. We note here that the indirect limits from BBN apply to SGWBs present in the early universe at the time of BBN (and characterized by an $\alpha = 0$ power law; see Eq 2), but not to SGWBs of astrophysical origin created more recently (and assumed to be characterized by an $\alpha = 3$ power law). Thus, the results presented here complement the indirect bound from BBN, which is only sensitive to cosmological SGWBs from the early universe, as well as direct $\alpha = 0$ measurements using lower-frequency observation bands [28].

There are several ways in which the methods presented in this paper can be improved. We list some ideas below:

(i) As mentioned in Sec. IVB, we can improve the estimate of the PEM contribution to the coherence by allowing for correlations between different PEM channels $\tilde{z}_i$ and $\tilde{z}_f$. This requires inverting the full matrix of PEM coherences $\gamma_{ij}(f)$ or solving a large number of simultaneous equations involving $\gamma_{ij}(f)$, rather than simply taking the maximum of the product of the coherences as was done here. A computationally cheaper alternative might be to invert a sub-matrix formed from the largest PEM contributors—i.e., those PEM channels that contribute the most to the coherence.

(ii) We can use bicoherence techniques to account for (non-linear) up-conversion processes missed by standard coherenc calculations. This may also allow us to identify cases where low-frequency disturbances excite higher-frequency modes in the detector.

(iii) The estimators $\hat{\Omega}_a(f)$ used in this analysis are optimal in the absence of correlated noise. In the presence of correlated noise, these estimators are biased, with expected values given by the sum, $\hat{\Omega}_a + \eta_a(f)$, where the second term involves the cross-spectrum, $N_{12}(f)$, of the noise contribution to the detector output. An alternative approach is to start with a likelihood function for the detector output $\hat{s}_1, \hat{s}_2$, where we allow (at the outset) for the presence of cross-correlated noise. (This should show up in the covariance matrix for a multivariate Gaussian distribution.) We can parametrize $N_{12}(f)$ in terms of its amplitude, spectral index, etc., and then construct posterior distributions for these parameters along with the amplitude and spectral index of the stochastic GW signal. In this (Bayesian) approach, the cross-correlated noise is treated on the same footing as the stochastic GW and is estimated (via its posterior distribution) as part of the analysis [58]. However, as described in [59], this works only for those cases where the spectral shapes of the noise and signal are different from one another.
GWs are produced by an unknown 'stiff' energy [54]. For the above figure, slow-roll inflationary model [52] assumes a tensor-to-scalar ratio of $r = 0.2$, the best fit value from the BICEP2 analysis [13]. In the axion based inflationary model, for certain ranges of parameters the backreaction during the final stages of inflation is expected to produce strong GWs at high frequencies [53]. The stiff equation of state (EOS) limit corresponds to scenarios in the early universe (prior to BBN) in which GWs are produced by an unknown 'stiff' energy [54]. For the above figure we used the equation of state parameter $w = 0.6$ in stiff EOS model. The cosmic string model corresponds to GWs produced by cosmic strings in the early universe [55]. The Earth's normal mode limits are based on the observed fluctuations in the amplitudes of Earth's normal modes using an array of seismometers [56]. The astrophysical SGWBs (BBH and BNS) are due to the superposition of Earth's normal modes using an array of seismometers [56]. The measurements of CMB and matter power spectra provide a similar integral bound in the frequency range of $10^{-15} - 10^{10}$ Hz [50]. The pulsar limit is a bound on the $\Omega_{GW}(f)$ at $f = 2.8$ nHz and is based on the fluctuations in the pulse arrival times from millisecond pulsars [51]. In the above figure, slow-roll inflationary model [52] assumes a tensor-to-scalar ratio of $r = 0.2$, the best fit value from the BICEP2 analysis [13]. In the axion based inflationary model, for certain ranges of parameters the backreaction during the final stages of inflation is expected to produce strong GWs at high frequencies [53]. The stiff equation of state (EOS) limit corresponds to scenarios in the early universe (prior to BBN) in which GWs are produced by an unknown 'stiff' energy [54]. For the above figure we used the equation of state parameter $w = 0.6$ in stiff EOS model. The cosmic string model corresponds to GWs produced by cosmic strings in the early universe [55]. The Earth's normal mode limits are based on the observed fluctuations in the amplitudes of Earth's normal modes using an array of seismometers [56]. The astrophysical SGWBs (BBH and BNS) are due to the superposition of coalescence GW signals from a large number of binary black holes (BBH) and binary neutron stars (BNS) [57].

(iv) We can also reduce correlated noise by first removing as much noise as possible from the output of the individual detectors. Wiener filtering techniques can be applied to remove acoustic, magnetic, and gravity-gradient noise from the time-series output of the LIGO detectors [60,62]. Furthermore, feed-forward control can be used to to cancel seismically-induced motion before it affects the LIGO test masses [61].

These and/or other techniques might be needed for future cross-correlation searches using advanced detectors, where improved (single-detector) sensitivity will mean that correlated noise may be an issue even for physically-separated detectors, such as the LIGO Hanford-LIGO Livingston detector pair [29,30,63].

Acknowledgment

The authors gratefully acknowledge the support of the United States National Science Foundation for the construction and operation of the LIGO Laboratory and the Science and Technology Facilities Council of the United Kingdom, the Max-Planck-Society, and the State of Niedersachsen/Germany for support of the construction and operation of the GEO600 detector. The authors also gratefully acknowledge the support of the research by these agencies and by the Australian Research Council, the International Science Linkages program of the Commonwealth of Australia, the Council of Scientific and Industrial Research of India, the Istituto Nazionale di Fisica Nucleare of Italy, the Spanish Ministerio de Educacion y Ciencia, the Conselleria d’Economia, Hisenda i Innovacio of the Govern de les Illes Balears, the Royal Society, the Scottish Funding Council, the Scottish Universities Physics Alliance, The National Aeronautics and Space Administration, the Carnegie Trust, the Leverhulme Trust, the David and Lucile Packard Foundation, the Research Corporation, and the Alfred P. Sloan Foundation.
