PAPER • OPEN ACCESS

Automated evaluation of environmental coupling for Advanced LIGO gravitational wave detections

View the article online for updates and enhancements.

Automated evaluation of environmental coupling for Advanced LIGO gravitational wave detections

A F Helmling-Cornell[*] [ID]**, P Nguyen**[1] [ID]**, R M S Schofield and R Frey** [ID]

University of Oregon, Eugene, OR 97403, United States of America

E-mail: ahelmlin@uoregon.edu

CrossMark

Abstract

The extreme sensitivity required for direct observation of gravitational waves by the Advanced LIGO detectors means that environmental noise is increasingly likely to contaminate Advanced LIGO gravitational wave signals if left unaddressed. Consequently, environmental monitoring efforts have been undertaken and novel noise mitigation techniques have been developed which have reduced environmental coupling and made it possible to analyze environmental artifacts with potential to affect the 90 gravitational wave events detected from 2015–2020 by the Advanced LIGO detectors. So far, there is no evidence for environmental contamination in gravitational wave detections. However, automated, rapid ways to monitor and assess the degree of environmental coupling between gravitational wave detectors and their surroundings are needed as the rate of detections continues to increase. We introduce a computational tool, PEMCHECK, for quantifying the degree of environmental coupling present in gravitational wave signals using data from the extant collection of environmental monitoring sensors at each detector. We study its performance when applied to 79 gravitational waves detected in LIGO's third observing run and test its performance in the case of extreme environmental contamination of gravitational wave data. We find that PEMCHECK's automated analysis identifies only a small number of gravitational waves that merit further study by environmental noise experts due to possible contamination, a substantial

---

[1] Now at: ZAP Energy, Everett, WA 98203, United States of America.
[*] Author to whom any correspondence should be addressed.

improvement over the manual vetting that occurred for every gravitational wave candidate in the first two observing runs. Building on a first attempt at automating environmental coupling assessments used in the third observing run, this tool represents an improvement in accuracy and interpretability of coupling assessments, reducing the time needed to validate gravitational wave candidates. With the validation provided herein; PEMCHECK will play a critical role in event validation during LIGO's fourth observing run as an integral part of the data quality report produced for each gravitational wave candidate.

1. Introduction

The era of gravitational wave (GW) astronomy began in 2015, when the Advanced LIGO (aLIGO) detectors directly observed the GWs from a binary black hole (BBH) merger [1]. In 2017, the first GWs from a binary neutron star (BNS) merger were observed by the aLIGO detectors [2]. Subsequent sky localization of GW170817's host galaxy was aided by the GW observatory Advanced Virgo, which did not observe GWs from the merger due to the orientation of the detector relative to the source [2–4]. In total the aLIGO detectors, along with Advanced Virgo and KAGRA, have observed 90 GWs from compact binary coalescences (CBCs) in the course of the first three observing runs [5–9].

The aLIGO detectors are a pair of nearly-identical observatories. LIGO Hanford Observatory (LHO) is located near Hanford, Washington, United States and LIGO Livingston Observatory (LLO) is located near Livingston, Louisiana, United States. Each detector is a kilometer-scale dual-recycled Fabry-Pérot Michelson interferometer [10, 11]. A GW incident on one of the aLIGO detectors produces an optical path length difference between light circulating in the two arm cavities of the interferometer. This varying difference in optical path length creates a time-dependent interference pattern when the light from the interferometer arms are recombined.

The amplitude of GWs is expressed as the GW strain h, where

$$h = \Delta L/L \tag{1}$$

with ΔL the differential arm length (DARM) between the two interferometer arms and L the unperturbed arm length, i.e. 4 km. Typical GWs observed by the aLIGO detectors have strain amplitudes $\mathcal{O}(10^{-21})$. This extreme sensitivity results in local environmental noise easily coupling to the detector output. Examples of environmental noise seen in the third observing run (O3)—which lasted from 1 April 2019–27 March 2020 with a commissioning break during the month of October 2019—include ground motion induced by air handling motors at LHO, ground motion from trains traveling near LLO and heightened anthropogenic activity at both sites during the normal workday increasing each detector's local ground motion [6, 7, 12–14]. Correlated noise between the two aLIGO detectors originating, for example, from lightning strokes, could potentially mimic or contaminate a GW signal [15, 16]. Given the precision needed to measure GWs and the diversity of potential environmental noise sources, it is crucial to identify times where the GW data may be contaminated by environmental effects.

When a GW is observed by one or more of the aLIGO detectors, detector characterization specialists must examine both the DARM data as well as data from a network of environmental sensors at each of the aLIGO sites and rapidly determine: (i) whether the GW candidate was a result of environmental effects in the DARM data, or (ii) whether there is environmental noise that is contaminating the signal, and if so, at what times and frequencies the noise

occurs [12, 17]. While there is no evidence so far that environmental noise has measurably affected a GW detection, the margin for safety grows smaller as the sensitivities of the aLIGO detectors improve.

In this paper, we report on a tool, PEMCHECK, developed to rapidly identify times and frequencies where environmental signals may couple to the DARM measurement made by the aLIGO detectors. The tool estimates the degree of environmental coupling and produces a recommendation for detector characterization experts to accept or reject a GW candidate event due to the absence or presence of environmental noise contamination in the GW data. This determination is typically completed in under 10 min.

This work represents a step forward in automated GW event validation. Prior to O3, any assessment of environmental contamination in a GW candidate was carried out entirely by hand [1, 18, 19]. During O3, a first attempt at automatically highlighting potential environmental contributions to GW signals was added as part of the data quality report (DQR) which was generated for each potential GW candidate [12, 17, 20, 21]. Its implementation led to a number of false positives where potential environmental coupling was identified by human event validators but was subsequently ruled out after physical environmental monitoring (PEM) experts reviewed PEM sensor data and found the previous version had often been overestimating the degree of environmental coupling. Transients in PEM channels near the GW candidates often showed that the claimed PEM couplings were inconsistent with the actual level of coupling during the event, usually at a frequency where only an upper limit estimate of the coupling was found (see section 2.2). In this work, we have improved upon this initial design in several ways. We have designed a metric, the contamination statistic, which aLIGO event validators can use to quickly establish whether there is substantial environmental noise in a GW candidate signal. We have also restricted automated coupling estimates to times and frequencies near GW candidates; previously the analysis was restricted only in time. In addition, we have implemented a sanity check for coupling estimates through a tuning procedure where we identify and reduce coupling factors which overestimate coupling to the GW channel. The inclusion of PEMCHECK in the fourth observing run (O4) DQR means that the environmental coupling assessment is available within minutes of a candidate event for the hundreds of candidate events expected to be observed in O4 [22–24].

This paper is organized as follows. In section 2 we briefly review the aLIGO PEM system and methods used to quantify coupling between PEM sensors and GW data. In section 3 we detail how PEMCHECK estimates the environmental coupling around GW candidates as well as how the contamination statistic is calculated. In section 4, we apply PEMCHECK to both simulated data and data from O3 GWs. In section 5 we outline future directions for improving the environmental coupling estimation method presented here.

2. aLIGO environmental monitoring

2.1. PEM system description

There are around 100 instruments at each aLIGO detector which monitor the detector's local environment. This PEM system is composed of accelerometers, microphones, magnetometers, voltage monitors, weather stations and other sensors. A map of all PEM sensors at each aLIGO detector is available at the aLIGO PEM group web page [25]. A thorough description of the PEM sensor network in its O3 configuration is found in [12]. Most PEM sensor data used for this work is sampled at 4096 Hz, although there are a few accelerometers and magnetometers which are sampled at 8192 Hz and 16 384 Hz. Detailed information on specific sensors and their configuration may be found in [12].

The aLIGO detectors have been modified in advance of the O4 to accommodate frequency-dependent quantum squeezing to improve the sensitivity of the observatories [26]. In addition, test masses at each detector were replaced due to defects in their mirror coatings [27]. The input laser power was increased. Several baffles were added to reduce scattered light noise and some existing baffles were damped to further reduce scattered light noise [14]. The septum window separating the output mode cleaner (OMC) from the rest of the interferometer was removed due to scattered light concerns as well. Prior to the start of O4, changes were also made to the PEM system. To reduce low-frequency quantum radiation pressure noise introduced by the addition of frequency-independent quantum squeezing in O3 [11], a $\sim$300 m long filter cavity (FC) was constructed at each aLIGO detector prior to O4 in order to realize frequency-dependent squeezing [26]. Accelerometers and magnetometers were added to the newly-constructed FC end station to monitor new potential noise coupling sites. A magnetometer monitoring the magnetic field near the LLO end-Y station (EY) seismic isolation controllers was added [28]. Large wire coils were installed to generate magnetic fields around the detectors to study the degree of magnetic coupling. The suite of microphones at LHO was replaced with upgraded hardware [29]. At both detectors, more computing space was allocated for PEM data storage so that GW signals up to 4 kHz could be vetted by PEM accelerometers, magnetometers and voltage monitors.

2.2. Studying environmental coupling

PEM sensor data can be used to determine coupling functions that quantify the degree to which environmental effects contribute to DARM. Coupling functions are determined by an extensive injection campaign prior to, after, and in the case of magnetometers, during each observing run. Detailed descriptions of injection procedures are found in [12, 30]. Here we summarize the process for determining the coupling function for a PEM sensor. A more thorough discussion is found in [12]. As an example, we will consider L1:PEM-CS_ACC_HAM6_OMC_Z_DQ, a channel associated with an accelerometer at LLO which records vacuum chamber motion in the vertical direction. This particular sensor is mounted atop horizontal access module (HAM) 6, the vacuum enclosure which contains the OMC—the meter-scale optical cavity which rejects unwanted light from the recombined signal from the interferometer arms, like modulation sidebands and higher order transverse modes [31]. The photodiodes used to measure DARM are also situated in HAM 6 where they witness the transmitted laser light leaving the OMC. Ground motion couples to DARM measurements by shaking the vacuum chamber walls or components on the optics table within this vacuum chamber, causing noise when light scattered from these shaking objects recombines with the main beam [14].

To quantify the response of the aLIGO detectors to environmental perturbations, coupling functions (CFs) between individual PEM sensors and the DARM data are produced. To measure a CF for a particular sensor, we perform many environmental noise injections across different frequency regimes and compare the response of the PEM sensor to the response in DARM. In the case of the OMC accelerometer, these include playing tones on a nearby speaker and mechanically shaking the nearby vacuum enclosure. To compute the CF for this OMC accelerometer, we compare the frequency-domain response of the PEM sensor and DARM during the injection. Specifically, we compute the amplitude spectral density (ASDs) of the accelerometer and DARM, denoted by $X_{\text{inj}}(f)$ and $Y_{\text{inj}}(f)$, respectively. These are compared to ASDs of sensor and DARM data from a quiet reference time where no environmental noise injection is occurring (denoted $X_{\text{bkgd}}(f)$ and $Y_{\text{bkgd}}(f)$, respectively).

At some frequencies, the detector couples strongly to the environment and the injected time DARM ASD clearly differs from the reference time ASD. For frequencies where this is the case, we compute the CF by [12]

$$M(f) = \sqrt{\frac{(Y_{\text{inj}}(f))^2 - (Y_{\text{bkgd}}(f))^2}{(X_{\text{inj}}(f))^2 - (X_{\text{bkgd}}(f))^2}}.$$
(2)

Here $M(f)$ indicates that these are 'measured' points since the detector response to an injection can be directly quantified by comparing the change in DARM and sensor ASDs.

At frequencies where the interferometers are well-isolated from environmental noise sources, a CF cannot be measured since the DARM ASD is unperturbed by the injection. In this case, we may only set an upper limit on the coupling between the PEM sensor and DARM. An upper limit estimate, $U(f)$, of the coupling may be set by assuming that at these frequencies, any noise in the DARM ASD is due to noise witnessed by the PEM sensor, or [12]

$$U(f) = \frac{Y_{\text{bkgd}}(f)}{\sqrt{(X_{\text{inj}}(f))^2 - (X_{\text{bkgd}}(f))^2}}.$$
(3)

The measured points and upper limit estimates on the coupling calculated via equations (2) and (3) are combined across several injections in different frequency bands to form a CF for each PEM sensor that runs from a few Hz to a few kHz. Each sensor's CF is a mix of measured points and upper limit coupling estimates. The value of the CF at a given frequency is chosen by identifying the highest-amplitude injection in the PEM sensor at that frequency (i.e. the injection with the largest value for $X_{\text{inj}}(f)$ at that frequency). The CF at the particular frequency takes the value of the measured or upper limit estimate of the coupling computed by equation (2) or (3) for this highest-amplitude injection. Figure 1 shows the CF computed for L1:PEM-CS_ACC_HAM6_OMC_Z_DQ by combining several PEM injections which was used during O3.

Potential sources of uncertainty for measured points in the CF arise from, for example, the finite spacing of PEM sensors. An environmental signal may be louder at an unmonitored coupling site than at the nearest sensor, resulting in an underestimation of the signal's effect on the GW data. Occasional environmental disturbances, such as the thunderclap discussed in section 4.2, show that predictions for DARM made using PEM and CF data correspond to the recorded value of DARM during the environmental transient with a factor of ∼2 uncertainty [12]. The true coupling between a PEM sensor and the DARM is given by a log-normal probability distribution since systematic uncertainties in the injection procedure prevent an entirely accurate measurement of the coupling at a given frequency [20].

We may estimate the contribution of environmental noise to the DARM measurement by projecting

$$Y_{\text{noise}}(f) = \text{CF}(f) \times X_{\text{GW}}(f)$$
(4)

where $X_{\text{GW}}(f)$ is the ASD of time series data in a given PEM sensor around the time of a GW candidate, $Y_{\text{noise}}(f)$ is the ASD of the DARM time series recorded at the GW candidate time solely due to environmental influences witnessed by the PEM sensor, and $\text{CF}(f)$ the sensor's CF.

For this work, we use CFs measured for O3 [34, 35]. Injection campaigns at both aLIGO detectors have been completed to update the CFs for their use in O4 [36, 37].
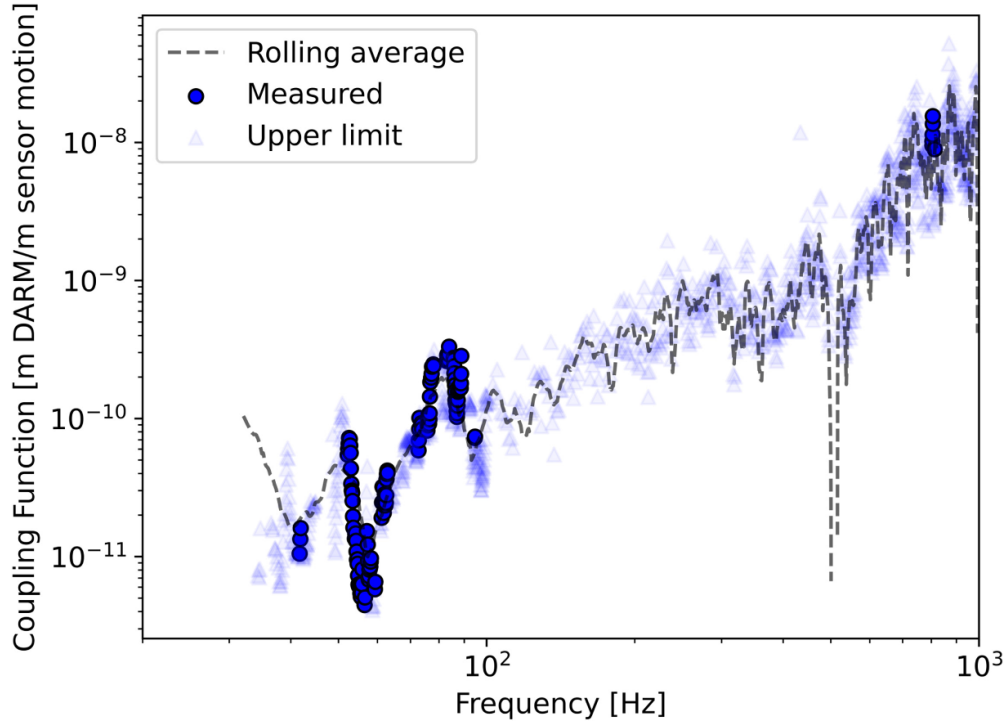
Figure 1. Vibrational coupling between LLO's HAM 6 vacuum chamber *Z*-axis accelerometer and DARM as measured prior to O3. The vibrational coupling here is likely driven by stray light scattering off of the septum window dividing this vacuum chamber and the adjoining HAM 5 chamber and then recombining with the main beam [32]. The large fraction of upper limit estimates, rather than measured values, for the vibrational coupling is because PEM injections could not be increased to an amplitude such that the injection was visible in the DARM data without saturating the accelerometer signal. We directly convert the acceleration measured by the accelerometer into displacement of the vacuum chamber wall. The resonance near 100 Hz could be a resonance of the vacuum chamber wall, the table, or optics on the table. Near 500 Hz the CF is set to 0 m/m because the detector is insensitive around the frequencies of the test mass suspension resonances (violin modes) [33].

3. Quantifying environmental coupling

3.1. Data access

The PEMCHECK analysis requires basic information about the GW candidate it is studying. This information can either be input manually by the user or it can be supplied automatically in the form of a GRACEDB Superevent ID [38]. If given a Superevent ID, PEMCHECK identifies the preferred event if there are a collection of triggers from different search pipelines all associated with the same GW candidate. We use the GWPY software package for data access and signal processing [39].

PEMCHECK is designed for data quality assessments of short-duration GW transients. While it is best-suited for studying GWs from CBCs, it can also be extended to analyze brief bursts of GWs from poorly modeled sources (e.g. cosmic strings, highly eccentric BBH mergers, etc) [40, 41]. For events identified by the low-latency CBC search pipelines GSTLAL, MBTA,

Table 1. Basic detection information and initial estimates of source frame masses and dimensionless spins for GW190707. This information is used to estimate the GW waveform shown in figure 4.

Online aLIGO detectors	LHO, LLO
t_c (GPS time, s)	1246 527 224.18
$m_1/M_\odot$	15.51
$m_2/M_\odot$	8.62
χ_1	-0.30
χ_2	0.76

PYCBC LIVE and SPIIR, PEMCHECK requires the masses and spins of each of the compact objects in the CBC as well as the GW candidate time [42–48]. The waveform parameters can be extracted from the GW candidate's preferred event data on GRACEDB. These parameters are used to approximate the time–frequency evolution of the GW signal in the aLIGO detectors using a non-precessing waveform model for CBCs which incorporates terms up to order 4 in the post-Newtonian expansion of the effective-one-body gravitational potential. The particular numerical relativity (NR) approximant used by PEMCHECK also employs reduced-order modeling of the potential which reduces the time needed to compute the GW waveform with a minimal effect on waveform accuracy [49]. In the case of a short, un-modeled GW transient, PEMCHECK requires the time and frequency ranges at which GW strain was observed by the cWB pipeline [50, 51]. This information is also available on GRACEDB if a cWB trigger is identified as the preferred event. The time and frequency information restricts the search for environmental coupling to the time and frequency ranges at which the GW candidate is witnessed by the detectors, which was not the case for the pre-O4 tool used to evaluate the presence of environmental noise in the DARM data. In addition to automatic parsing of GRACEDB data, PEMCHECK also accepts manual entry of waveform parameters or rectangular regions in time–frequency space where excess energy is observed in the GW data. In the case of a noteworthy GW candidate during O4, such as the first detection of GWs from a new source type, PEM experts will still vet the data by hand.

Here we use GWTC-3 event GW190707_09 3326 (hereafter GW190707) as an example to describe how PEMCHECK estimates the extent to which environmental noise contaminates the GW data [52, 53]. Table 1 lists the inferred merger properties used by PEMCHECK for waveform approximation. We restrict the following discussion to LLO data.

3.1.1. DARM and PEM data.   Once the time and frequency ranges of the GW candidate are established, PEMCHECK requests DARM and PEM sensor data from around the candidate time. The DARM data is sampled at 16 384 Hz. This time series data is split into foreground and background times. For CBCs, the default definition for the foreground time is $t_\text{fore} = [t_{20}, t_c]$, where t_{20} is the earliest time at which the GW signal is greater than or equal to 20 Hz and t_c is the merger time as specified by the GRACEDB preferred event. For each candidate event, a suitable background time period is also identified. For a CBC signal, the minimum background time is $t_\text{back} = [t_{20} - 32\,\text{s}, t_{20})$. For a burst signal, $t_\text{fore} = [t_\text{start}, t_\text{end}]$ and $t_\text{back} = [t_\text{start} - 32\,\text{s}, t_\text{start})$, with t_start and t_end the beginning and end of the time period identified by the burst search pipeline. The foreground time is defined to capture the behavior of the PEM sensors and the GW data during the GW candidate itself. The background time is chosen to provide the longer-term behavior of these channels. A longer background time means that any glitches in the GW channel which occur during the background are averaged over, reducing
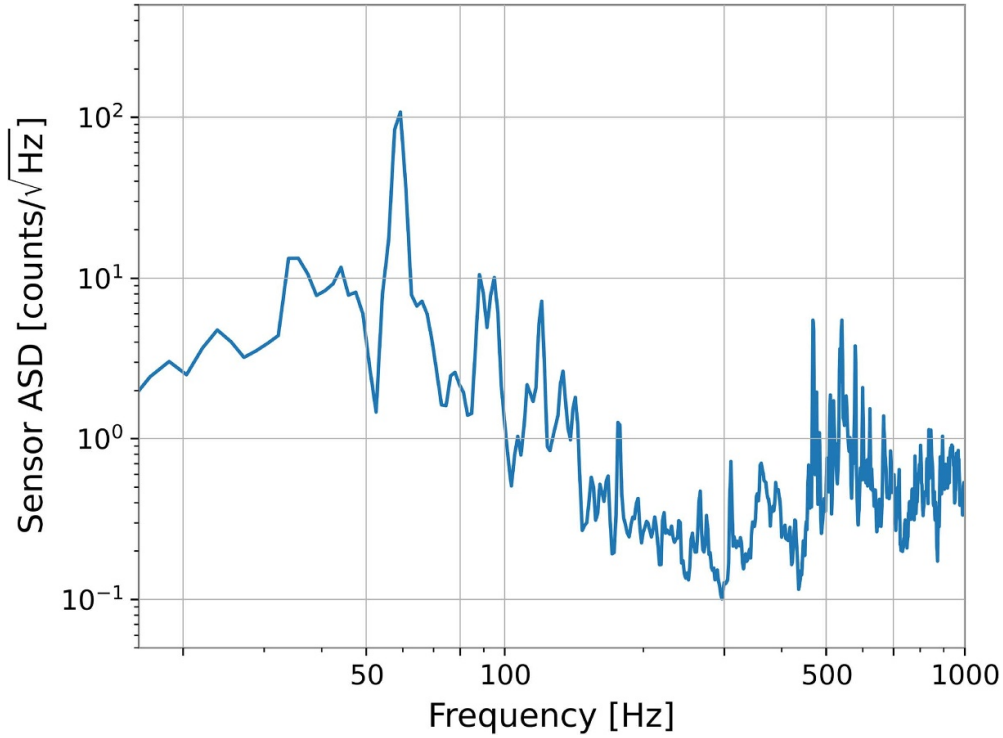
Figure 2. ASD of the Z-axis HAM 6 accelerometer during the background time. Background PEM sensor data is used to compute the tuned CF of section 3.2.1. 45.75 s of sensor data were used to calculate this ASD. Note that each sensor count corresponds to 6.1 μm s^{-2} of acceleration in the vertical direction [25].

the amount they affect this calculation. In the case of a longer-duration GW signal, the background time can extend to up to 256 s prior to t_{20} or t_{start} for a CBC or burst signal, respectively. The background duration is increased from the minimum when the duration of the foreground period is comparable to the minimum background duration of 32 s. In the case of GW190707, $t_{20} = 1\,246\,527\,219.46$ s and t_c is given in table 1.

3.2. ASD projection

Using the time series data described in section 3.1.1, we calculate the ASD of DARM and each PEM sensor data during both the foreground and background time segments. The duration, τ, of each fast Fourier transform used in the ASD calculation is given by

$$\tau = \begin{cases} 0.25\,\text{s} & \text{if } t_{\text{fore}} < 2\,\text{s} \\ t_{\text{fore}}/8 & \text{if } 2\,\text{s} \leqslant t_{\text{fore}} \leqslant 8\,\text{s} \\ 1\,\text{s} & \text{otherwise} \end{cases} \tag{5}$$

and the overlap fraction is $\max(1/2, 1 - t_{\text{fore}}/16\tau)$. Welch's method with median averaging and a Hann window is used to calculate the ASD for each channel [54]. Figure 2 shows the ASD of the PEM sensor background data.

We now discuss the projection of environmental noise into the DARM data during t_{fore}. For each PEM sensor studied by PEMCHECK, we first linearly interpolate between the CF data points so that the frequency resolution of the CF matches the frequency resolution of the GW data. We then tune the CF for the sensor if the predicted level of noise in DARM due to the PEM data exceeds the actual level of noise during the background time. We next compute the ASD of the PEM sensor and DARM data using the prescription given in equation (5). Then, for each frequency bin in the PEM ASD, we compute the meters of DARM during the GW candidate caused by environmental noise via equation (4).

3.2.1. CF tuning.   Sometimes the CFs used for PEMCHECK overestimate the level of environmental noise present in the GW data. To reduce environmental noise coupling overestimation, we *tune* the interpolated CFs supplied to PEMCHECK for each GW candidate. This process works by comparing $Y_{\text{noise}}(f)$ during the background time to the actual background time DARM ASD. At frequencies where the environmental noise ASD is predicted to exceed the DARM ASD, we reduce $CF(f)$ at those frequencies such that the predicted environmental noise is equal to the value of the DARM ASD at those frequencies. This process is illustrated in the top plot in figure 3. By reducing the CF at points where there is no evidence for the claimed coupling during t_{back} we limit the instances of overestimating environmental coupling.

Overestimating environmental coupling may lead to event validators spending additional time investigating noise sources identified by PEMCHECK that are not truly present in the GW data. Typically this manual review process involves identifying transients in the PEM sensor data at the same frequency as the reported coupling and close to the GW candidate in time and checking whether the predictions for environmental noise contamination are consistent for these loud transients. There may also be difficulties interpreting the astrophysical parameters of a given GW event. The difficulty disentangling the effects of offline noise subtraction from novel features in a GW signal was highlighted in the context of noise subtraction due to glitches—not environmental noise—during O3 [55–57]. Event validators incorrectly prescribing noise subtraction due to PEMCHECK overestimating the degree of environmental noise present in a GW signal may lead to similar parameter estimation issues in O4.

In the case of the LLO HAM 6 accelerometer and GW190707, the CF overestimates the level of environmental noise contribution to the DARM measurement, particularly between $\sim 30 - 40\,\text{Hz}$ and near 90 Hz, as illustrated in both plots of figure 3. In this case, environmental coupling projections derived from upper limit estimates of the coupling, rather than measured points, in the CF exceed the actual level of noise in the detector data. The coupling estimate is tuned at these frequencies to be consistent with the observed GW data. The tuning procedure reduces the need for manual vetting of GW signals by down-weighting spurious predictions like this one.

Disagreement between the HAM 6 accelerometer CF and the observed coupling could arise due to differences between the ASD used in calculating the reference ASD during sub–100 Hz vibrational and acoustic injections and the background ASD data shown in figure 2 [12]. These differences may arise due to different noise sources being present in the data in March and April 2019, when the CFs were measured and July 2019, when that GW was observed. Additionally, commissioning efforts in the intervening months increased the low–frequency sensitivity of the LLO detector, further constraining upper–limit estimates of the environmental coupling where it was not measured.

By tuning each CF to account for the environmental noise and condition of the aLIGO detector at the GW candidate time, we improve the accuracy of environmental noise
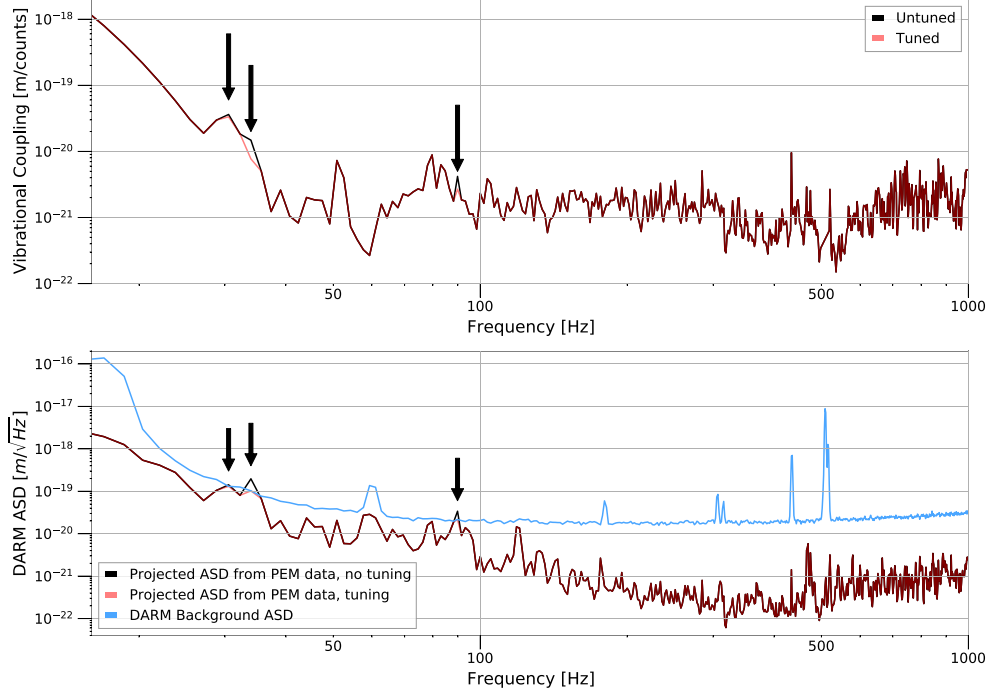
Figure 3. Top: Comparison of the CF interpolated from the data in figure 1 to the interpolated, tuned CF computed by analyzing the DARM background ASD. The untuned and tuned predictions for the ASD largely agree, except near 30 and 90 Hz. The points where environmental contributions to DARM are overestimated and need tuning are derived from upper limit estimates of the chamber motion to DARM coupling. These frequencies are marked with black arrows. Bottom: comparison of the DARM t_{back} ASD with the estimated contribution from the HAM 6 accelerometer. The result of the tuning procedure is to constrain the predicted environmental contribution from a sensor to be not greater than the observed DARM ASD. As in the top plot, black arrows mark where the untuned CF predicts that the DARM ASD should exceed its observed value to the HAM 6 accelerometer data. These are the points which require tuning to reconcile the predicted and actual DARM ASDs.

contamination information presented to aLIGO event validators compared to the information given to O3 event validation experts [12, 17].

3.3. c-statistic determination

From the ASD of the DARM data during the candidate GW event as well as the predicted DARM ASD curve due to environmental noise witnessed by a particular PEM sensor, we quantify the degree to which environmental disturbances may contaminate GW strain data.

Following equation (1), we divide the tuned CF by a factor of 4000 m (the length of the aLIGO Fabry–Pérot arm cavities) to convert the CF from units of DARM per sensor unit observed to units of GW strain per sensor unit [10]. We then compute the spectrogram of the GW strain data during the foreground time as well as the spectrogram of the effective strain measurement due to environmental coupling for each PEM channel, as described in section 3.2.

We denote the predicted value of the GW strain spectrogram tiles due solely to environmental contamination as $\mu(t,f)$, where

$$\mu(t,f) = (\mathrm{CF}_{\mathrm{tuned}}(f)/L) \times S_{\mathrm{PEM}}(t,f) \tag{6}$$

and $S_{\mathrm{PEM}}(t,f)$ is the foreground-time spectrogram of the PEM sensor data. The predicted spectrogram is then compared to the spectrogram of the GW strain, which we denote as $h(t,f)$.

For frequencies where μ was computed using a measured CF, then—as mentioned in section 2.2—the of observing a strain h due to a local environmental disturbance is given by the log-normal probability density function with a standard deviation of $\ln(2)$, or

$$\rho(h,\mu) = \frac{1}{h(t,f)\ln(2)\sqrt{2\pi}} \exp\left(-\frac{\left(\ln(h(t,f)) - (\ln(\mu(t,f))\right)^2}{2(\ln(2))^2}\right). \tag{7}$$

For frequencies where μ was computed using an upper-limit CF, we use a uniform probability distribution on the range $[0,\mu)$ to estimate the probability density function of observing the GW strain, i.e.

$$\rho(h,\mu) = \begin{cases} \frac{1}{\mu(t,f)} & \text{if } h(t,f) < \mu(t,f) \\ 0 & \text{otherwise} \end{cases}. \tag{8}$$

The strict limit on the probability density function here is because the pre-observing run PEM injections are unable to establish a coupling between the sensor data and the GW data at these frequencies. For GW strains that are larger than the stringent upper limits set on the coupling times the PEM sensor data, we assume that the likelihood of environmental noise coupling vanishes.

To describe the likelihood of environmental coupling near a GW event we define the contamination statistic, denoted by c, which quantifies the likelihood of recording a strain value in excess of $h(t,f)$ due to environmental noise alone. We compute it via the cumulative distribution function (CDF) of the distributions in equations (7) and (8). The c-statistic at each point in the GW spectrogram is given by

$$c(t,f) = \begin{cases} \frac{1}{2}\left[-\mathrm{erf}\left(\frac{\ln(h(t,f)) - \ln(\mu(t,f))}{\sqrt{2}\ln(2)}\right) - 1\right] & f \in (f_{\mathrm{meas}}) \\ \min\left(\frac{h(t,f)}{\mu(t,f)}, 1\right) & f \notin (f_{\mathrm{meas}}) \end{cases}. \tag{9}$$

where f_{meas} denotes the set of frequencies at which the CF was measured, rather than an upper limit on the CF estimated. Here erf() is the error function and min() denotes the minimum of the two numbers.

The c-statistic is distributed over the range $[0,1]$. A low c-statistic for a given sensor indicates that environmental disturbances witnessed by the PEM sensor could account for some of the data in the GW strain channel at a given time and frequency range. A high c-statistic indicates that it is unlikely that environmental noise witnessed by the PEM sensor is coupling to the GW data. Environmental contamination of the GW data is indicated by one or more sensors reporting low c-value, inconsistent with the low-noise case, where many sensors should report a c-statistic near 1. Results from computing the c-statistic many times over all the PEM data channels are shown in section 4.1. These distributions largely reflect the expectation for the distribution of the c-statistic for the null hypothesis (negligible environmental contamination), as discussed in section 4.1.

Not every spectrogram tile's c-statistic is relevant to determining whether there is environmental noise contamination in the GW strain signal. For instance, aLIGO detected GWs at a maximum frequency of $\sim$693 Hz during GW190707; any environmental noise that was present
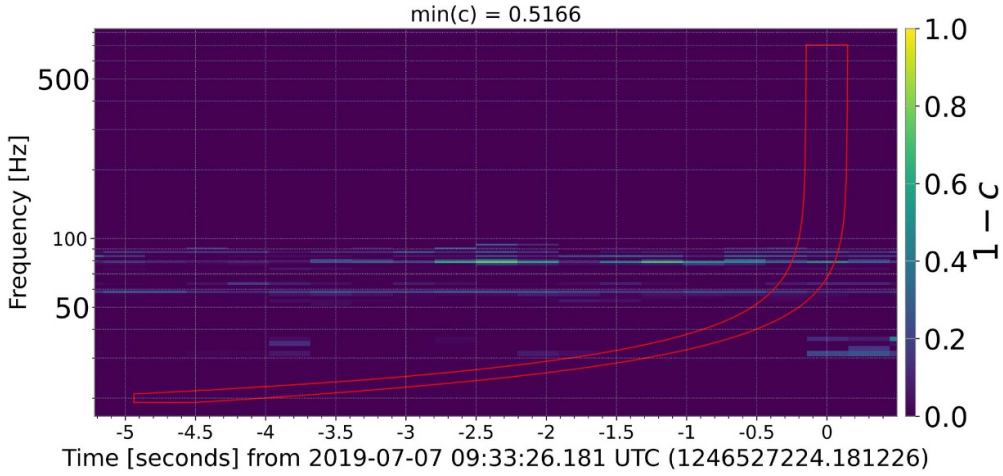
Figure 4. The results of the c-statistic calculation for each spectrogram tile for the LLO HAM 6 accelerometer during GW190707. The value of $1 - c$ is plotted for each tile so that time–frequency tiles with the highest likelihood of contamination, as calculated by PEMCHECK, appear brighter. The time–frequency tiles enclosed by the red box track the evolution of the GW signal as predicted by NR. The lowest c-statistic found within the GW track occurs at $\sim 0.1\,$s and $\sim 78\,$Hz prior to the event's t_c. A minimum c-statistic of 0.52 indicates that it is unlikely that environmental disturbances witnessed by this sensor couple to the GW strain data during the event.

in kHz frequencies, whether it couples strongly or not to the DARM measurement, should not affect the GW signal reconstruction. Therefore, we only consider the c-statistic for spectrogram tiles whose central frequencies are within 1–4 Hz of the time–frequency track predicted by the NR waveform approximant or the time–frequency 'box' predicted by the GW burst search. The results of the c-statistic calculation for the LLO accelerometer during GW190707 are shown in figure 4.

3.4. Combining results

PEMCHECK reports to the user the time–frequency region with the lowest c-statistic from the entire collection of PEM channels analyzed. From the results presented in section 4, we find that a GW candidate with a minimum c-statistic of 0.2 or less for one or more PEM channels seems to be a reasonable threshold in O4 for follow up by environmental noise experts. This threshold identifies GW candidates with the most significant evidence for environmental coupling while minimizing the number of events needed to be hand-vetted. If this threshold is too restrictive in O4, it may be changed. For each GW, a report may be automatically generated showing the time and frequency location of the minimum c-statistic tile near the GW data for each PEM sensor.

The coupling estimation process outlined here is valid only for linear coupling between one or more PEM channels and the GW data. Analyzing nonlinear coupling—where environmental disturbances at a particular frequency pollute GW data at a different frequency—is typically done from a glitch mitigation perspective, and is beyond the capability of PEMCHECK at this time.
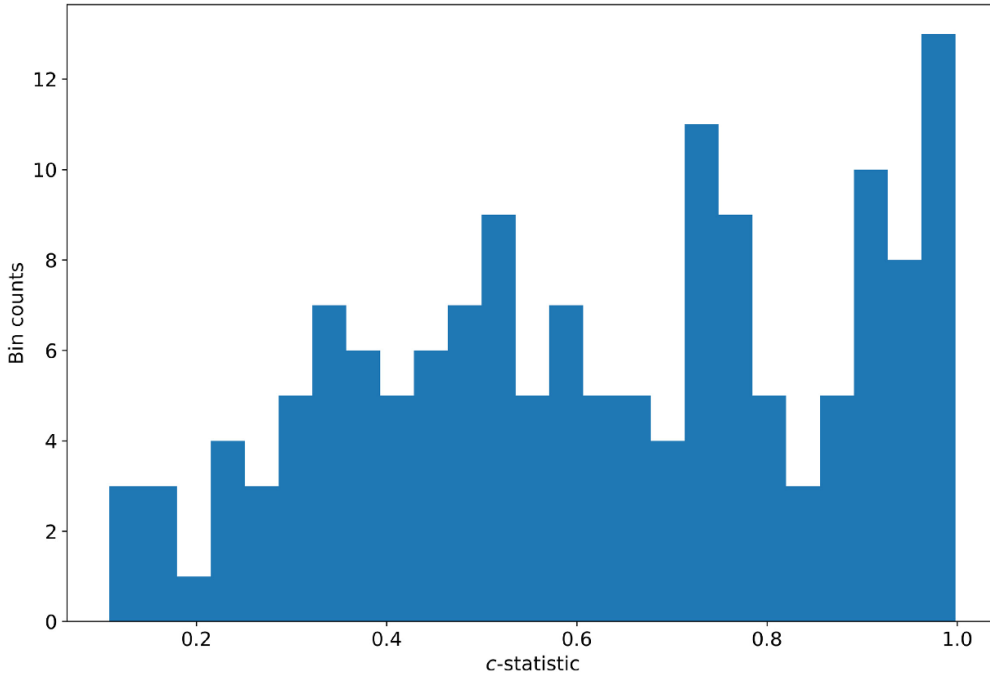
Figure 5. Histogram of the minimum c-statistic calculated by PEMCHECK for all O3 events using the pre-O3 coupling data. The PEMCHECK analysis identifies 6 of the 149 O3 GW events as having a minimum c below 0.2.

4. Results

4.1. Real GW data

We report the results of running PEMCHECK on data from each aLIGO detector around all 79 GW events detected by aLIGO in O3 with a of astrophysical origin of 0.5 or greater [6, 7, 33, 58]. This amounts to 149 individual runs of PEMCHECK, one per GW event per online aLIGO detector. The sensor CFs determined prior to the start of O3 were used for the following analyses, with the exception of results presented in table 2.

We find that PEMCHECK does not report a c-statistic of 0.1 or less for any of the 149 runs. The distribution of the minimum c-statistic for all 149 instances of PEMCHECK is shown in figure 5. Figure 6 shows the cumulative distribution of c-statistics for all channels for all O3 events analyzed in this work. Overall, most GW events report a large c, or a low likelihood of environmental noise contamination. However, there are a few outliers in this set which merit further consideration. While all the steps involved with manually vetting GW candidates is beyond the scope of PEMCHECK, it provides a starting point for where to look for evidence of environmental coupling. In sections 4.1.1–4.1.3 we discuss the events with the lowest minimum c-statistics in more detail.

We also find that the tuning procedure was frequently used to down-weight the predicted level of environmental contamination. Table 2 shows the number of events where the time–frequency tile with the highest likelihood of contaminating the GW data was derived from a tuned projection of the PEM channel coupling. The data indicate that the CFs found via the PEM injection procedure predicted realistic estimates of environmental noise about two-thirds
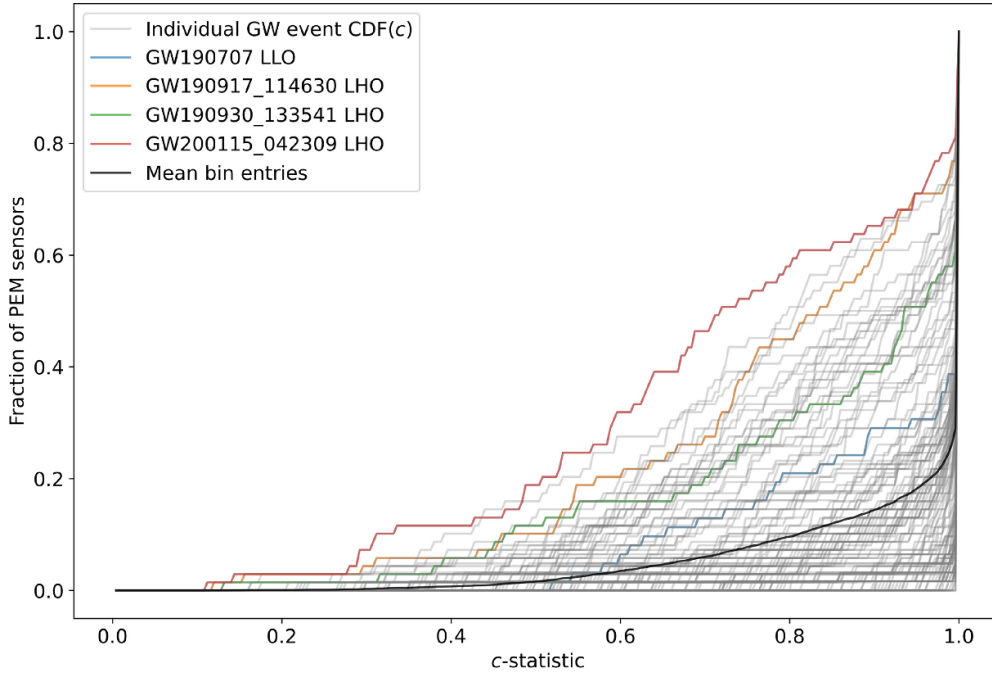
Figure 6. Cumulative distribution of the c-statistic found for each PEM channel studied by PEMcheck for each GW event in O3 using the pre-O3 coupling data. The coupling analysis was performed on 63 PEM sensors at LLO and 72 PEM sensors at LHO. Each light gray trace corresponds to the CDF of the c-statistic for an individual GW event, while the black trace denotes the mean number of sensors with that particular value of c or less. GW190707 as well as the three events with the lowest c are highlighted.

Table 2. The number of O3 GW events where the tuning procedure was used in calculating the time–frequency tile with the highest likelihood of contamination. The type of point, measured or upper limit, in the worst coupling sensor CF at that frequency is also given. Most, but not all, tuning is done at frequencies where sensor CFs are given by upper limit estimates of the coupling. We compare the fraction of tuned, measured points identified by PEMcheck when the CFs measured prior to O3 are used to the fraction of tuned, measured points identified when the CF data for LHO's squeezer table accelerometer is replaced with the CF data collected after O3 for GW events occurring after 13 August 2019.

	Untuned		Tuned	
CF data used	Measured	Upper limit	Measured	Upper limit
Pre-O3	61	35	22	31
Pre-O3 except for LHO squeezer table	65	43	7	34

of the time. However, about one-third of events had initial noise projections which were inconsistent with the actual DARM ASD and needed to be tuned. In many of these cases an upper limit estimate of the coupling, where no response to a PEM injection is seen in the GW data, turns out to overestimate the environmental noise. This is the intended utility of the tuning procedure. However there were also 22 instances of a measured coupling between a PEM

sensor and the GW channel overestimating the environmental noise contamination. Of these, 16 stem from a measured coupling at either 148 or 196 Hz between an accelerometer, H1:PEM-CS_ACC_ISCT6_SQZLASER_X_DQ, on the optics table which houses many of the components of LHO's squeezer. This accelerometer's coupling consistently needed to be tuned down at 148 Hz starting with GW190814, which was observed the day after the squeezer table configuration was modified due to the original squeezer laser failing [59, 60]. It is possible that the squeezer configuration changes affected the coupling between squeezer optics table motion and the GW channel. After reanalyzing the dataset with this sensor's CF as measured at the conclusion of O3, we find that only 7 minimum c-statistic tiles are derived from tuned, measured values of the coupling. None of the remaining 7 tuned, measured points stem from the LHO squeezer table vibrational coupling being overestimated. For these 7 points, tuning reduced the predicted environmental noise in the GW channel by a median of 18% compared to the untuned prediction for the GW channel ASD at the worst coupling frequency.

The tuning procedure, which was designed primarily to provide better estimates of the coupling between a sensor and the GW data where the coupling could not be measured due to PEM sensor saturations, works as intended for the O3 events once the LHO squeezer accelerometer correction is applied. Event validation experts using PEMCHECK no longer have to investigate unrealistically large claims of environmental effects in the GW data. Instead, this is now done automatically by the tuning process. In addition, tracking which measured values of the pre-run CFs consistently need tuning during events may indicate to detector commissioners that changes to the instrument made during the run may have affected the environmental coupling of the detector and that some subset of the CFs must be re-measured.

4.1.1. GW200115_042309.   The PEMCHECK analysis of data near GW200115_042309 (hereafter GW200115) in LHO produced the lowest single c-statistic of the entire dataset. This GW event, first reported in [61], is likely the result of a neutron star-black hole collision.

The time–frequency tile with the lowest c-statistic is located 0.144 s prior to the merger near 1417 Hz. The PEM channel which produced the lowest c-statistic tile is H1:PEM-CS_MAG_LVEA_VERTEX_XYZ_DQ with $c = 0.108$. This channel is not recorded in aLIGO GW frame files but is instead calculated by the PEMCHECK code by taking the quadrature sum of three channels, each corresponding to a different axis of the triaxial magnetometer placed near the LHO beamsplitter which directs input light down the two interferometer arms.

We manually investigate magnetometer and GW data near the time–frequency tile identified by PEMCHECK to evaluate the claim of environmental coupling. Comparing the constant-Q transforms of the strain and magnetometer channels, we find that there is some signal at that time and frequency in the magnetometer data [62–64]. However, there are other instances of magnetic signals at that frequency near the GW time which are more energetic which do not produce any response in the DARM channel. For this reason, we do not conclude that there is environmental contamination from the local magnetic environment in this instance. The difficulty associated with obtaining a measured magnetic coupling at high frequencies led to the environmental coupling being overestimated. The tuning procedure was not initiated in this instance; the upper limit estimate on the coupling in the region of 1417 Hz is characterized by a comb feature where the upper limit estimate of magnetic field to DARM coupling is considerably larger in a few Hz band. Above 1417 Hz, the coupling is tuned down, but this frequency happens to be the point in the comb that, when the projected ASD is calculated, comes the closest to the actual background DARM ASD without exceeding it. Since PEMCHECK's results are subject to human review, the apparent overestimation of magnetic coupling at this point would have been identified during the internal event validation process.
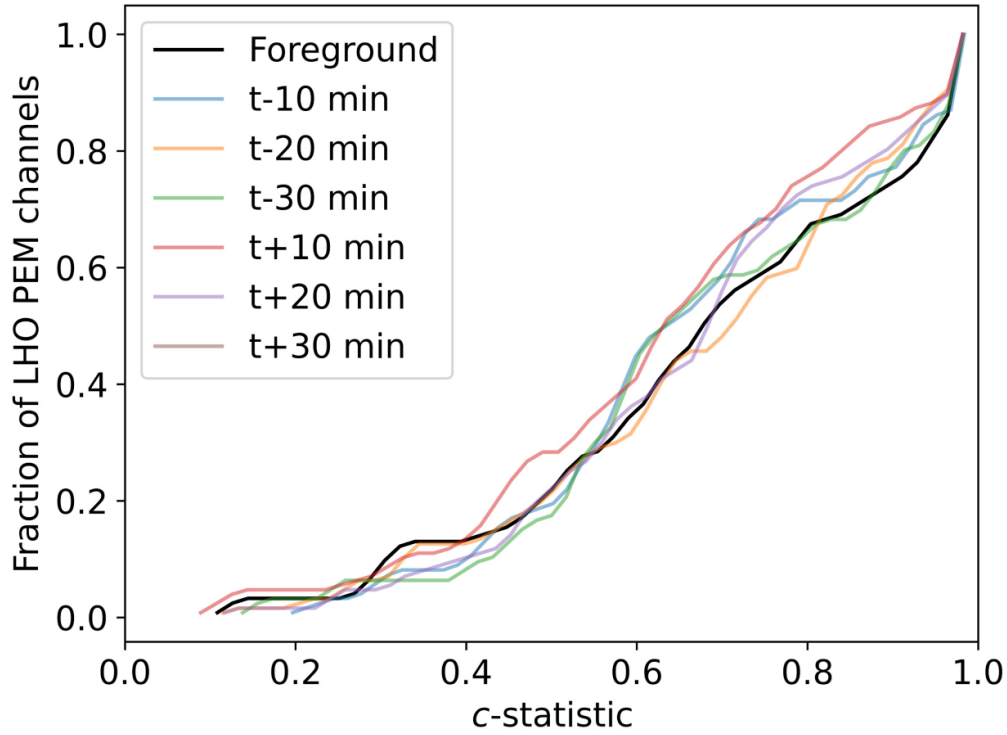
Figure 7. CDF of the *c*-statistic for each run of PEMCHECK at LHO with GW200115's properties. The CDF marked 'foreground' corresponds to PEMCHECK run at the actual time of the GW event.

GW200115 also stands out as having many channels report a moderate *c*-statistic. Its CDF(c) is almost always the highest of all the other traces shown in figure 6. Of the sensors identified by PEMCHECK with a $c \leqslant 0.7$, 10 corner station accelerometers report the highest evidence for coupling at 50–53 Hz at 3.64 s prior to the merger time. This is above the *c* threshold of 0.2 proposed as the value of the *c*-statistic needed to initiate manual review. To confirm that these channels are not coupling to DARM, we examine the distribution of *c*-statistics at nearby times to determine whether the collection of corner station (CS) accelerometers at near-threshold *c* is exceptional. As illustrated in figure 7, the distribution of channel *c*-statistics changes little when PEMCHECK is run using the same parameters inferred in low-latency for the time–frequency track at ±10, 20 and 30 min from the actual GW time. Like the foreground time, the distribution of moderately-valued *c*-statistics is in each instance of the PEMCHECK computation driven by corner station accelerometers. Ground motion at the corner station is heightened due to split mini air handling units operating in the time around the GW detection. No noise transients due to the air handling units are seen in the hour of strain data centered on the real GW event. The lack of evidence for environmental coupling from these channels indicates that a minimum *c* of 0.2 or less is likely a reasonable threshold for manual review.

4.1.2. GW190917_114630. The event with the second lowest value of min(c) is GW190917_114630 at LHO. PEMCHECK finds that the potential environmental contamination

may be witnessed by the pair of triaxial magnetometers located in the electronics room at end-X station (EX). The minimum c-statistic tile is at 837 Hz and 0.084 s after the merger time and has $c = 0.117$. Following the same argument as in section 4.1.1, the constant Q-transforms of magnetometer and GW data do not support environmental contamination in the data. Again, poorly-constrained high-frequency magnetic CFs are responsible for overestimating the potential for environmental contamination in this case.

4.1.3. GW190930_133541.   The final event with a low value of c we examine is GW190930_133541 at LHO. An LHO accelerometer monitoring motion of the vacuum enclosure which houses the signal recycling cavity at 50.6 Hz and 0.352 s prior to the merger time is responsible for the c-statistic of 0.129 for this GW event [25, 65]. This purported coupling is not supported by studying constant-Q transforms of the GW strain and PEM channel data near the event. The CF data for this sensor near the frequency identified by PEMCHECK is set by an upper limit estimate of the coupling, which is an order of magnitude estimate rather than a more precise measure of the coupling, as described in section 2.2. This could explain the discrepancy between the coupling projection and the lack of environmental coupling seen in the GW data from around the event. There is not support for environmental coupling for this event.

While the methodology of manual vetting of GW candidates is beyond the scope of this work, we present some arguments usually employed in the manual vetting process to establish that this event is free from environmental contamination from this PEM sensor. We compare constant Q-transforms of PEM sensor and GW data from the GW event time to look for occurrences of coincident high-energy transients in the two datastreams. These two Q-transforms are shown in figure 8. There are time–frequency tiles at 50.6 Hz in the accelerometer Q-transform with higher energy than the tile identified by PEMCHECK as having the lowest c-statistic. Some examples of these transients are found at approximately $t_c - 1.5\,\text{s}$ and $t_c + 1.1\,\text{s}$. There is no corresponding response to these transients in the GW data. These higher energy tiles were ignored by PEMCHECK because they lie outside the calculated time–frequency track of the GW signal. Second, we investigate the frequency-domain behavior of the signal recycling cavity beamtube accelerometer to see whether it is exceptional in some way near the time of GW190930_133541. Longer-duration spectrograms of channel data indicate that the accelerometer is not witnessing any extraordinary environmental noise at 50.6 Hz during the event time. Lastly, the estimated ambient noise level witnessed by this sensor at this frequency is close to the GW sensitivity when projected into GW data [12, 25]. Because of this, random fluctuations in the PEM ASD data should occasionally lead to claims of coupling that need to be followed up on.

4.2. Simulated GW signals

We examine the performance of PEMCHECK in an extreme scenario: one or more pipelines claim detection of a GW during a period of high environmental contamination of the DARM data. Thunderclaps can induce a signal in DARM by abruptly increasing ground motion in the tens of Hertz band [12, 17]. An especially loud thunderclap was visible in DARM as well as accelerometers located at LLO's EY station in May of 2019 [66]. We consider the results of running PEMCHECK on 37 hypothetical GW signals with the properties of GW190521_030229 (hereafter GW190521) occurring during this particular thunderclap. The properties of GW190521 used for this analysis are given in table 3.
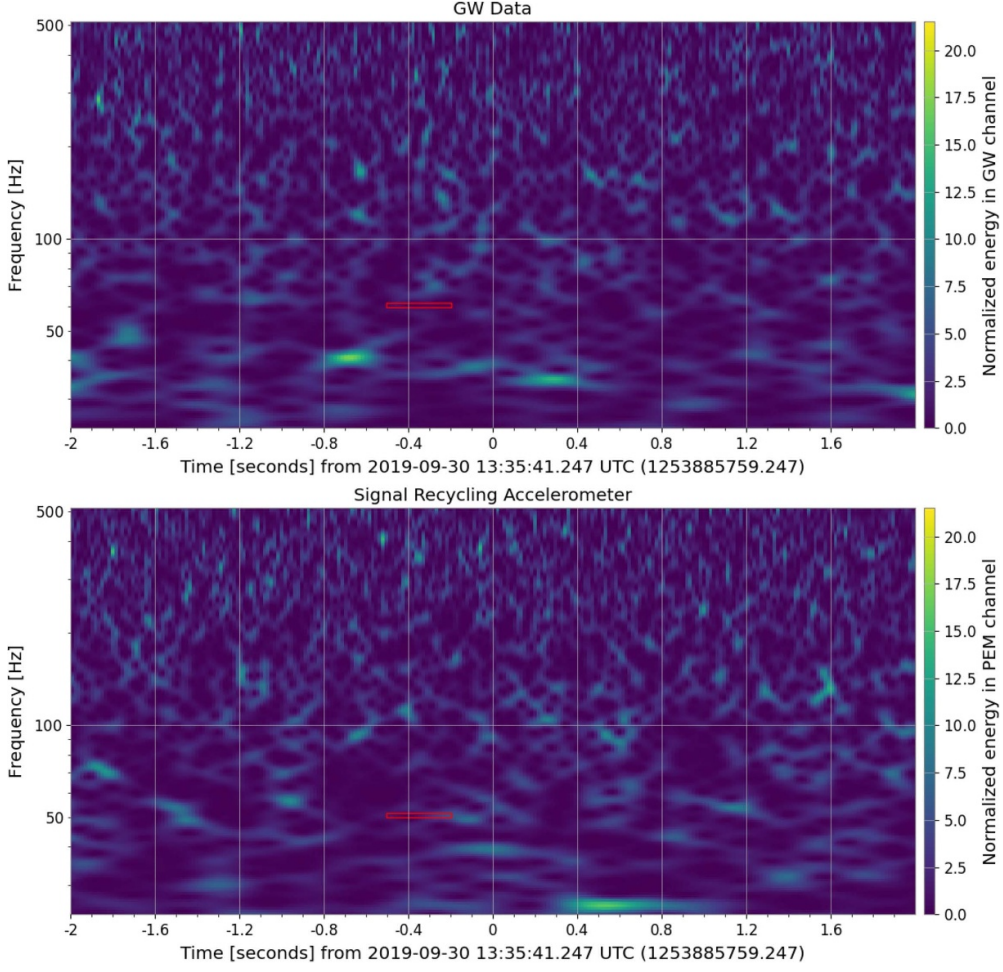
Figure 8. Constant-Q transform of LHO GW strain (top) and signal recycling cavity beamtube accelerometer (bottom) data centered on GW190930_133541's arrival time. A red box is overlaid on each plot showing the time–frequency window identified by PEMCHECK as having the highest likelihood of coupling. Higher-energy transients in the accelerometer compared to the tile flagged by PEMCHECK appear in the accelerometer data but do not correspond with transients at 50.6 Hz in the GW data. This indicates that CF is overestimating the coupling at this point and that the data is not subject to environmental contamination. Furthermore, the coupling estimate by PEMCHECK is driven by the transient beginning near $t_c - 0.2$ s, but the NR-approximated inspiral waveform lies at a higher frequency by the time this transient appears.

We find that PEMCHECK routinely identifies a low c in the fictitious GW signals during the thunderclap shown in figure 9. The frequency identified as the likeliest one to be contaminated is 48 ± 2 Hz in each trial. The tiles with the most normalized energy in the 40–50 Hz band correspond to GW strains of 2.54×10^{-22}. During the thunderclap, the L1:PEM_EY_ACC_BSC5_ETMY_Y_DQ accelerometer recorded a maximum displacement of 1.22×10^{-8} m. Projecting this into the GW strain signal with the sensor's CF, we find that the predicted strain value associated with that amount of sensor motion at that frequency is

Table 3. Estimates of source frame masses and dimensionless spins for GW190521 taken from [67]. These parameters were used to generate the tracks in figure 9. No specific t_c is given, as it differs for each of the 37 GW tracks in the analysis.

Progenitor property	Value
$m_1/M_\odot$	85
$m_2/M_\odot$	66
χ_1	0.69
χ_2	0.73

Figure 9. Below: constant-Q transform of LLO GW strain data during a thunderclap witnessed by PEM sensors. The 37 red time–frequency tracks overlaid on the thunderclap data represent 37 hypothetical GW time–frequency tracks with the properties given in table 3 vetted by PEMCHECK. Above: the lowest value of c found by PEMCHECK for the series of hypothetical GW tracks overlaid on thunderclap data. The channel with the lowest c-statistic for a given trial is denoted by the marker style. The bold line corresponds to $c = 0.2$.

2.36×10^{-22}. As a point of comparison, GW150914's strain reached a maximum amplitude of 2.63×10^{-22} in the same frequency band [68].

Should an event validation expert see a GW candidate with multiple nearby sensors reporting a low c-statistic at similar frequencies, an appropriate recourse would be to contact environmental monitoring and noise subtraction experts for further analysis and study of the environmental noise background local to the candidate event.

5. Conclusions

In this work, we have demonstrated the feasibility of the PEMcheck tool to identify potential coupling between the aLIGO detectors and their environment for GW candidate events. PEMcheck uses CFs measured between PEM sensors and the DARM measurement to project noise witnessed by PEM sensors into the GW strain data stream and determine the likelihood that the GW data at a given time and frequency is a result of environmental noise. This is quantified by the contamination (or c) statistic.

We have tested the performance of PEMcheck by analyzing the 79 GW confident events seen by one or more of the aLIGO detectors in O3. This amounts to 149 individual runs of PEMcheck. The events with the lowest value of the contamination statistic are studied in further detail. When subject to manual inspection, we conclude that the potential environmental coupling is not borne out in each case. This is juxtaposed with a clear case of environmental coupling caused by a thunder clap. From the results of running PEMcheck on real and simulated GW signals a criterion of $c \leqslant 0.2$ is suggested as the standard for initiating human review of PEMcheck results by event validation experts.

We show that PEMcheck is capable of warning event validation experts if there is substantial environmental coupling. We find that the tuning procedure used by PEMcheck provides a rough estimate of the environmental coupling in real-time, reducing the demand for hand-vetting of events by environmental noise experts. However, claims of environmental coupling continue to be checked by hand in the case of a high likelihood of environmental contamination or detections of GWs from novel sources. PEMcheck is being used in the DQR to study environmental noise contamination in every GW observed by the aLIGO detectors in O4.

Future directions for development may include closer integration with utilities for characterizing changes in sensor performance over time [69] and designing a framework for constraining coupling from PEM sensors using loud transients seen in those channels and not in DARM to further tune estimates of the CF, similar to the approach demonstrated in section 4.1.3 partially manually vetting GW190930_133541. The results of the tuning procedure should be tracked to motivate injected campaigns if CFs tend to become inaccurate as observing runs continue, especially in cases where upper limit estimates consistently need adjusting. Nonlinear coupling is not accounted for by current CFs and thus PEMcheck, but would be an important development for the future. A more rigorous approach to trials factors may improve the interpretability of numerical results. The limitations of this procedure, e.g. in PEM sensor coverage, will inform the designs for future GW detectors, such as Einstein Telescope and Cosmic Explorer, for which environmental noise will continue to be crucial to confront [70, 71]. Since the PEMcheck analysis requires CFs, sensor and DARM data, the same procedure used to search for evidence of environmental coupling could be used in these future GW detectors.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://pem.ligo.org/.

ORCID iDs

A F Helmling-Cornell ● https://orcid.org/0000-0002-7709-8638
P Nguyen ● https://orcid.org/0000-0001-9223-5541
R Frey ● https://orcid.org/0000-0003-0341-2636

References

[1] Abbott B P *et al* (LIGO Scientific Collaboration and Virgo Collaboration) 2016 *Phys. Rev. Lett.* **116** 061102
[2] Abbott B P *et al* (LIGO Scientific Collaboration and Virgo Collaboration) 2017 *Phys. Rev. Lett.* **119** 161101
[3] Abbott B P *et al* (SKA South Africa/MeerKAT) 2017 *Astrophys. J. Lett.* **848** L12
[4] Acernese F *et al* 2015 *Class. Quantum Grav.* **32** 024001
[5] Abbott B P *et al* (LIGO Scientific Collaboration and Virgo Collaboration) 2019 *Phys. Rev.* X **9** 031040
[6] Abbott R *et al* 2021 arXiv:2108.01045
[7] Abbott R *et al* (LIGO Scientific Collaboration, Virgo Collaboration, and KAGRA Collaboration) 2023 *Phys. Rev.* X **13** 041039
[8] Acernese F *et al* (Virgo Collaboration) 2019 *Phys. Rev. Lett.* **123** 231108
[9] Akutsu T *et al* 2021 *Prog. Theor. Exp. Phys.* **2021** 05A101
[10] Aasi J *et al* 2015 *Class. Quantum Grav.* **32** 074001
[11] Buikema A *et al* 2020 *Phys. Rev.* D **102** 062003
[12] Nguyen P *et al* 2021 *Class. Quantum Grav.* **38** 145001
[13] Glanzer J, Soni S, Spoon J, Effler A and González G 2023 *Class. Quantum Grav.* **40** 195015
[14] Soni S *et al* 2021 *Class. Quantum Grav.* **38** 025016
[15] Janssens K *et al* 2023 *Phys. Rev.* D **107** 022004
[16] Washimi T Yokozawa T, Nakano M, Tanaka T, Kaihotsu K, Mori Y and Narita T 2021 *J. Instrum.* **16** 07033
[17] Davis D *et al* 2021 *Class. Quantum Grav.* **38** 135014

[18] Abbott B P *et al* (LIGO Scientific Collaboration and Virgo Collaboration) 2016 *Class. Quantum Grav.* **33** 134001
[19] Schofield R *et al* 2017 H1 L1 PEM vetting report for the Aug. 17, BNS candidate (available at: https://dcc.ligo.org/LIGO-T1800508/public)
[20] Nguyen P 2022 Environmental noise in gravitational wave detectors and the search for gravitational wave signals associated with gamma-ray burst during LIGO'S third observing run *PhD Thesis* University of Oregon
[21] The LIGO Scientific Collaboration and The Virgo Collaboration 2018 Data quality report user documentation (available at: https://docs.ligo.org/detchar/data-quality-report/)
[22] The LIGO-Virgo-KAGRA Collaboration 2022 Dqrtasks user documentation (available at: https://detchar.docs.ligo.org/dqrtasks/)
[23] Abbott R *et al* 2023 *Phys. Rev.* X **13** 011048
[24] Abbott B P *et al* (KAGRA Collaboration, LIGO Scientific Collaboration and Virgo Collaboration) 2018 *Living Rev. Relativ.* **21** 3
[25] Schofield R M S PEM central (available at: https://pem.ligo.org/) (Accessed 2023)
[26] Ganapathy D *et al* (LIGO O4 Detector Collaboration) 2023 *Phys. Rev.* X **13** 041021
[27] Brooks A F *et al* 2021 *Appl. Opt.* **60** 4047–63
[28] Malaquias-Reis J A and Effler A 2023 Pem sensor work at ey (available at: https://alog.ligo-la.caltech.edu/aLOG/index.php?callRep=64021)
[29] Clara F and Helmling-Cornell A 2023 Pem microphone upgrade (available at: https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=69448)
[30] Effler A, Schofield R M S, Frolov V V, González G, Kawabe K, Smith J R, Birch J and McCarthy R 2015 *Class. Quantum Grav.* **32** 035017
[31] Arai K, Barnum S, Fritschel P, Lewis J and Waldman S 2013 Output mode cleaner design (available at: https://dcc.ligo.org/LIGO-T1000276/public)
[32] Schofield R 2022 Ham5/6 vibration coupling much reduced after septum window removal, 35.5 hz darm peak is from residual coupling in the area (available at: https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=66470)
[33] Abbott R *et al* 2023 *Astrophys. J. Suppl. Ser.* **267** 29
[34] Nguyen P *et al* 2019 Pre-o3 pem injection summary and links to individual logs (available at: https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=49521)
[35] Schofield R, Austin C P N and Effler A 2019 Llo magnetic injections -o3 couplings (available at: https://alog.ligo-la.caltech.edu/aLOG/index.php?callRep=45371)
[36] Short R, Glanzer J, Helmling-Cornell A, Effler A and Schofield R 2023 Pem injections successfully completed - will inform hvac experiments (available at: https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=69745)
[37] Effler A *et al* 2023 Pem injections status (available at: https://alog.ligo-la.caltech.edu/aLOG/index.php?callRep=64609)
[38] LIGO Scientific Collaboration Gracedb | the gravitational-wave candidate event database (available at: https://gracedb.ligo.org/) (Accessed 2023)
[39] Macleod D, Areeda J S, Coughlin S B, Massinger T J and Urban A L 2021 *SoftwareX* **13** 100657
[40] Abbott R *et al* (LIGO Scientific Collaboration, Virgo Collaboration, and KAGRA Collaboration) 2021 *Phys. Rev. Lett.* **126** 241102
[41] Abac A G *et al* 2023 arXiv:2308.03822
[42] Messick C *et al* 2017 *Phys. Rev.* D **95** 042001
[43] Sachdev S *et al* 2019 arXiv:0912.1470
[44] Adams T, Buskulic D, Germain V, Guidi G M, Marion F, Montani M, Mours B, Piergiovanni F and Wang G 2016 *Class. Quantum Grav.* **33** 175012
[45] Nitz A H, Dal Canton T, Davis D and Reyes S 2018 *Phys. Rev.* D **98** 024050
[46] Dal Canton T and Harry I W 2017 arXiv:1705.01845
[47] Hooper S, Chung S K, Luan J, Blair D, Chen Y and Wen L 2012 *Phys. Rev.* D **86** 024012
[48] Chu Q 2017 Low-latency detection and localization of gravitational waves from compact binary coalescences *PhD Thesis* Western Australia University
[49] Bohé A *et al* 2017 *Phys. Rev.* D **95** 044028
[50] Klimenko S, Mohanty S, Rakhmanov M and Mitselmakher G 2005 *Phys. Rev.* D **72** 122002
[51] Klimenko S *et al* 2016 *Phys. Rev.* D **93** 042004
[52] Abbott R *et al* 2021 *Phys. Rev.* X **11** 021053

[53] LIGO Scientific Collaboration and Virgo Collaboration 2019 GCN 25012 (available at: https://gcn.gsfc.nasa.gov/other/GW190707q.gcn3)
[54] Welch P 1967 *IEEE Trans. Audio Electroacoust.* **15** 70–73
[55] Huber M C and Davis D 2022 arXiv:2208.13844
[56] Hannam M *et al* 2022 *Nature* **610** 652–5
[57] Payne E Hourihane S, Golomb J, Udall R, Davis D and Chatziioannou K 2022 *Phys. Rev. D* **106** 104017
[58] Farr W M, Gair J R, Mandel I and Cutler C 2015 *Phys. Rev. D* **91** 023005
[59] Abbott R *et al* (LIGO Scientific Collaboration and Virgo Collaboration) 2020 *Astrophys. J. Lett.* **896** L44
[60] Dwyer M, Kijbunchoo N, Merfeld K, Sigg D and McCarthy R 2019 Comment to "replacing the failing squeezer laser (available at: https://alog.ligo-wa.caltech.edu/aLOG/index.php?callRep=51244)
[61] Abbott R *et al* (the LIGO Scientific Collaboration, the Virgo Collaboration, and the KAGRA Collaboration) 2021 *Astrophys. J. Lett.* **915** L5
[62] Chatterji S, Blackburn L, Martin G and Katsavounidis E 2004 *Class. Quantum Grav.* **21** S1809–18
[63] Chatterji S 2005 The search for gravitational wave bursts in data from the second LIGO science run *PhD Thesis* Massachusetts Institute of Technology
[64] Brown J C 1991 *J. Acoust. Soc. Am.* **89** 425–34
[65] Arain M A and Mueller G 2008 *Opt. Express* **16** 10018
[66] Nguyen P and Schofield R 2019 Coupling estimates from pem injections correctly predict llo darm glitch near s190510g from thunder-driven vibration at ey (available at: https://alog.ligo-la.caltech.edu/aLOG/index.php?callRep=46025)
[67] Abbott R *et al* (LIGO Scientific Collaboration and Virgo Collaboration) 2020 *Phys. Rev. Lett.* **125** 101102
[68] Abbott R *et al* (The LIGO Scientific Collaboration The Virgo Collaboration) 2021 *SoftwareX* **13** 100658
[69] The LIGO Scientific Collaboration and The Virgo Collaboration 2018 ligocam user documentation (available at: https://docs.ligo.org/detchar/data-quality-report/tasks/ligocam-dqr.html)
[70] Punturo M *et al* 2010 *Class. Quantum Grav.* **27** 194002
[71] Reitze D *et al* 2019 Cosmic Explorer: the U.S. contribution to gravitational-wave astronomy beyond LIGO *Bull. Am. Astron. Soc.* **51** 35