Identifying multichannel coherent couplings and causal relationships in gravitational wave detectors

Piljong Jung,[1] Sang Hoon Oh,[1] Young-Min Kim,[2] Edwin J. Son,[1] Takaaki Yokozawa,[3] Tatsuki Washimi,[4] and John J. Oh[1,*]

[1]*Gravity Research and Application Team (GReAT), National Institute for Mathematical Sciences (NIMS), Daejeon 34047, Republic of Korea*
[2]*Department of Physics, Ulsan National Institute of Science and Technology (UNIST), Ulsan 44919, Republic of Korea*
[3]*Institute for Cosmic Ray Research (ICRR), KAGRA Observatory, The University of Tokyo, Kamioka-cho, Hida City, Gifu 506-1205, Japan*
[4]*Gravitational Wave Science Project (GWSP), Kamioka branch, National Astronomical Observatory of Japan (NAOJ), Mitaka City, Tokyo 181-8588, Japan*

The gravitational-wave detector is a complex and sensitive collection of advanced instruments that are impacted not only by mechanical/electronics systems but also by the surrounding environment. Hence, it is of great importance to classify and mitigate noises to detect gravitational-wave signals by using information from many auxiliary channels related to such devices and surroundings. This improves the signal-to-noise ratio and reduces false alarms from coincident loud events. For this reason, it is essential for identifying coherent relationships between complex channels. This study presents a way of identifying (non)linear couplings between associated channels by using the method of correlation coefficients. We show that the method can be applied to practical problems in the gravitational-wave detector, such as noises by lightning strokes, air compressors vibrations, and noises caused by wind effects.

I. INTRODUCTION

The detection of a gravitational wave (GW) emitted from the binary black hole coalescence (GW150914) [1] opened the era of GW astronomy, which is expected to pursue new perspectives of understanding the structure of the Universe and the evolution of astrophysical objects such as stars and galaxies. For a more profound understanding and expansion of the knowledge of the Universe, there still exist many challenges to be overcome for GW physics and astronomy. Much more accurate sky localization for GW sources and a much farther range of luminosity distance in GW observation requires the improved sensitivity of GW detectors, overcoming the present limitation of technologies.

Detectors' sensitivity confines the boundary of observations so that the number of GW events within the boundary determines the detection rate that can be achieved in the current GW detector. The currently operating ground-based laser interferometric GW detectors such as Laser Interferometer Gravitational-wave Observatory (LIGO), Virgo, and the KAGRA have a similar design sensitivity curve with a frequency band of 30–2000 Hz, which is characterized by three primary noise sources: photon shot noise by the laser system in the high-frequency range, thermal noise by test-mass mirrors in the midfrequency range, and seismic noise by ground vibrations in the low-frequency range [2–5]. Those GW detectors are now being operated for detecting GW signals and planned for upgrades to improve their sensitivities using many engineering challenges such as cryogenics, quantum squeezed light, and so on. Furthermore, next-generation ground-based GW detectors with new conceptual designs are now being planned [6,7].

Besides, advanced methodologies and analysis algorithms should be required to enhance the quality of data [8–10] taken from the very sensitive instruments and isolate GW signals from noises caused by the sensitive devices and the surrounding environments efficiently because it improves the detection statistics such as signal-to-noise ratio (SNR) to provide the reliable detection criterion. This, consequently, also improves the detection range of GW detectors, which yields better detection rates of GW sources. In this point of view, both developments of advanced methodologies for data analysis and understanding the status of GW detectors, as well as the most sensitive instruments, are of special importance to achieve for detecting GW signals coming from the farthest GW emitting objects. For this reason, many different tools for characterizing noises have been developed and utilized for analyzing GW signals so far. The first purpose of those

---

*johnoh@nims.re.kr

tools is to categorize and classify transient/continuous noises that are harmful to the GW strain channel, identifying the causal relationship in the coincident families of noisy channels. Then they should be mitigated with the help of various advanced mathematical algorithms as possible. Finally, if they are caused by some instrumental defects with a possibility of repeated malfunctions, the causes should be reported and amended for maintaining the consistent status of the detector.

Studies of finding coherence between the strain channel and auxiliary channels of the GW detectors have been extensively performed in LIGO-Virgo collaborations, in the context of continuous GW and stochastic GW background searches [11–16]. Many efforts for identifying and vetoing transient noises have been made so far and are widely utilized for GW data analysis, such as a computation of significance called *hierarchical veto (Hveto)* [17] and *Used percentage veto (UPV)* [18], noises associated with long-duration transients [19,20], Q-transform based trigger generator [21], the Hilbert-Huang transform-based method [22], the bicoherence method [23], linear regressions [24], machine learning algorithms [25], and so on. These methods cover classification and a vetoing method as well as identification of the influences between the GW strain channel and auxiliary channels monitoring the environmental and/or instrumental status. This needs a consistent understanding of which channel effects can cause the transient noises in the GW strain channel and/or which can be useful to mitigate such abnormalities to maintain the normal status of the detector. However, we still face great challenges in dealing with the noises of GW data because the GW detector behaves with a highly nonstationary and nonlinear nature.

For this reason, it is necessary to develop more advanced analysis tools for noise hunting to improve the detector's data quality. Here we focus on the couplings between a certain auxiliary channel and the GW strain channel. In Ref. [26], it has been shown that the excess noises from bilinear and nonlinear couplings in GW interferometers can be treated using the bilinear coupling veto (BCV) method [27] with the trigger-based correlation coefficient. Along this line, we suggest a new way of identifying coherent associations between the GW strain channel and the related auxiliary channels, in which we use three kinds of the correlation coefficient: Pearson's correlation coefficient (PCC) [28] and Kendall's τ coefficient (Ktau) [29] as a linear measure and maximal information coefficient (MIC) as a nonlinear measure [30–32], respectively. In particular, MIC is an information-theoretic measure to discriminate the nonlinear association between two random variables. Together with these measures, we construct a systematic way of identifying the noises from (non)linear couplings and causalities propagating from instrumental and/or environmental disturbances of GW detectors. Then we apply the suggested method to the well-known issues of

data analysis and noise identification in the KAGRA detector, expecting that the correlation unidentified in the previous methods can be verified as a nonlinear correlation effect.

In this study, we present a way of identifying coherent associations between the GW strain channel and auxiliary noise channels by computing correlation measures. The method we consider here is the PCC and Ktau as a linear measure and MIC as a nonlinear measure. We construct a consistent way of discriminating the relevant noise effect and its causality, applying it to some issues in GW detection. In Sec. II, we describe the methods to measure (non)linear associations and build an analysis process with proper statistical algorithms between the GW strain channel and many auxiliary channels. In Sec. III, we exhibit the exemplary results based on some noise data taken from the KAGRA GW detector; a lightning stroke, air compressor noises, and the noises caused by the wind effect. Finally, we summarize and discuss our results in Sec. IV.

II. METHOD AND WORKFLOW

We use the data taken from the GW strain channel and many auxiliary channels in the KAGRA GW detector. The KAGRA is a gravitational-wave detector with a similar configuration of laser interferometry such as the LIGO/ Virgo except for the cryogenic test-mass mirrors and the underground installation [33]. The initial installation of the KAGRA has finished in 2019, and after year-long commissioning, it started its first observing run during a month in 2020 and joined the O3 observing run together with the advanced LIGO and the Advanced Virgo, recently [34].

The major feature of the KAGRA is the cryogenic and underground GW detector, which implies that the KAGRA detector has somewhat unique characteristics caused by the nature of the underground cryogenic facility, producing the relevant noise effects that were not reported in the ground-based GW detectors. Therefore, it is of great importance to understand the noise characteristic of the KAGRA detector and its environment.

A. Methods for data correlations

We introduce three methods of analysis for investigating the correlation between two data samples; the PCC, the Ktau, and the MIC. Let us consider two time-series data X and Y to be nonstationary and univariate datasets with an equal size of n. We assume that (x_i, y_i) is a set of ith bivariate data pairs from the paired data (X, Y). If we assume that the dataset Y includes certain noises, being affected by the sensitivity of GW detectors, there exists a meaningful statistical association between two observed variables X and Y because the noise can propagate to another dataset X. With these assumptions, we analyze the correlated relationship of noises resulting from the

instrumental anomalies and/or environmental interference in GW detectors. In this section, we describe three major methods to estimate the linear and nonlinear associations based on the time-series data recorded in each channel.

The PCC is defined as the ratio between the covariance and the product of standard deviation of each variable, which produces the linear correlation score. More precisely, let $\bar{x} = \sum_{i=1}^{n} x_i/n$ and $\bar{y} = \sum_{i=1}^{n} y_i/n$ be the means of X and Y, respectively, then, the PCC ρ is

$$\rho(X, Y) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}, \quad (1)$$

which determines the value between -1 and 1. For independent two variables, we have $|\rho| = 0$. When $|\rho| > 0.5$, it is typically interpreted as a significant correlation. $|\rho| = 1$ indicates that two variables have a perfect linear correlation, where positive and negative signs imply increasing and decreasing linear dependence, respectively. Here, we used the absolute value of the PCC for convenience.

The Ktau [29] measures the strength of the monotonicity of the relationship between two variables, which is defined by

$$\tau(X, Y) = \frac{2(c - d)}{n(n - 1)}, \quad (2)$$

where c and d are the number of concordant and discordant pairs in (X, Y), respectively. Given two data samples from the combined variable set, $(X, Y) = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$, if sampled pairs are either $x_i > x_j$ and $y_i > y_j$ or $x_i < x_j$ and $y_i < y_j$, it is called a concordant pair. On the other hand, if they are either $x_i > x_j$ and $y_i < y_j$ or $x_i < x_j$ and $y_i > y_j$, it is called a disconcordant pair. Hence, Ktau provides the ordinal association that is proportional to the difference between ordered and disordered pairs in all possible combinations. As with the PCC, $\tau(X, Y)$ varies from -1 to 1. If the order of two pairs is randomly distributed, they are monotonically independent, yielding $\tau(X, Y) = 0$. If Y values tend to change with increasing or decreasing X values, the absolute value of Ktau becomes one.

Meanwhile, the mutual information (MI) can estimate and characterize the strength of shared information between two random variables. For given two discrete random variables A and B with a joint probability mass function $p(a, b)$ and marginal probability mass functions $p(a)$ and $p(b)$, MI is defined by

$$I(A; B) = \sum_{a \in A} \sum_{b \in B} p(a, b) \log_2 \frac{p(a, b)}{p(a)p(b)}. \quad (3)$$

Note that MI provides non-negative values, $I(A; B) \geq 0$. If the random variable B is a function of A, $B = f(A)$, $I(A; B)$ diverges. In addition, if A and B have no shared information, $p(a, b) = p(a)p(b)$, then $I(A; B)$ clearly vanishes; they are statistically independent. Refer to [35] for more detailed properties of MI.

Suppose that the joint probability distribution $p_{XY}(i, j)$ is approximated by the number of points falling into the i-by-jth bin on the partitioned plane of scattered plots by X and Y. Then, the approximated MI is obtained by the occupied elements in each cell as

$$I(X; Y)_{k,l} = \sum_{i=1}^{k} \sum_{j=1}^{l} \frac{p_{XY}(i, j)}{\log_2 \min\{k, l\}} \log_2 \frac{p_{XY}(i, j)}{p_X(i)p_Y(j)}, \quad (4)$$

where $p_X(i)$ and $p_Y(j)$ are marginal distributions on the ith column and jth row. In addition, k and l indicate the partition size of the X column and Y row, respectively. Because the Jensen's inequality, $0 \leq I(X; Y)_{k,l} \leq \log_2 \min\{k, l\}$, is satisfied [35], Eq. (4) can be normalized and the value spans between zero and one.

To overcome a heuristic approach of maximizing the MI value for all possible resolutions of k-by-l grid [30], three intrinsic properties in Eq. (4) are utilized; they are (1) monotonic convex function, (2) bounded set, and (3) uniformly continuous function, which allows computing Eq. (4) more efficiently and effectively.

Let S be a n-ordered pair set and $B(n) = n^{\alpha}$ and the parameter c restricts the total grid size. Here, α is a dimensionless parameter that controls the size of grids and c is controlling the coarseness of maximizing the discrete grid search. Then, the empirical MIC called MICe is defined as

$$MICe(X, Y, \alpha, c) = \max_{ab < B(n)} \left\{ \frac{\max I^{[*]}(S, k, l)}{\log_2 \min\{k, l\}} \right\}, \quad (5)$$

where $I^{[*]}(S, k, l)$ denotes the maximized MI in $k \times l$ grids. Note that the coefficient of MICe varies from zero to one. If two time-series data have a clear association, Eq. (5) converges to one for sufficiently large data sizes whereas it becomes zero in the opposite condition. Even though MICe has a unique performance to extract the complex associations, it was uncertain how reliable MICe values can be provided in the aspect of the parameter selection. To guarantee the reliability of computed MICe, we suggested an empirical way of determining the relevant parameters of MICe through the statistical power to optimize them in [36].

B. CAGMon: A novel tool for identifying correlations

We utilize the aforementioned correlation measure indices to identify the data associations between the GW strain channel and auxiliary channels of GW detectors. To monitor the chronological trend of association coefficients between two datasets, we considered a minimal bin of data segments, called *stride*, which represents a piece of an equal interval unit in the whole data segments. For given datasets, the aforementioned three correlation coefficients, $\{\rho(X_m, Y_m), \tau(X_m, Y_m), MICe(X_m, Y_m)\}$, are computed

for divided subsets of X_m and Y_m by stride. By gathering the sequential coefficients together, we can demonstrate a time-serial trend of the association strength between two channels. In this section, we present a novel tool for identifying and diagnosing the association of two datasets, called *CAGMon*; the workflow and design of the algorithm.

The workflow of *CAGMon* tool is depicted in Fig. 1, which comprises four different stages: (i) reading initial configuration, (ii) loading time-series data and preprocessing, (iii) computing each coefficient and the relevant statistics, and (iv) plotting results and building a result summary page. First, the configuration file consists of user-defined parameters, preprocessing options, and general features of data such as sampling rate, start/end times, and stride. Because each coefficient is computed for each stride of time-series datasets, the tool can only investigate the similar timing coincidence within the same stride bin. Hence, there was a limitation in identifying the association

FIG. 1. A flowchart of the *CAGMon* tool.

between transient noises correlated to noise with a specific frequency. To overcome this limitation, we provide several customized options for data preprocessing, such as high-pass, low-pass, and bandpass filters, for a variety of scalable analyses. In addition, the general feature of data comprises a primary channel name, the path of data files, an auxiliary channel list file, segment file, output file type and its save path, and so on. When the time-series data is retrieved from GW frame files, the configuration option refers to the initial configuration file to apply the preprocessing. If time-series data contain either blank or NaN, zero arrays would be padded properly to avoid a computational error. Then, data are rescaled to have the same data size using the resampling algorithm. The detailed discussion and analyses on the reliability of MICe values, data sample size, and optimal parameters of MICe have been presented in [36]. In this study, we select a set of optimal parameters of MICe presented in Table I.

In Fig. 2, the *CAGMon* tool presents: (i) a correlation matrix of ρ, τ, and MIC between two datasets, (ii) a scattered pattern between both channels, and (iii) the correlation trend behavior during a certain time period between two channels. The correlation method simply produces the coherence between two datasets. So one requires a way of improving the statistical significance of a detected observation. Quantitatively, the way of using timeslide studies in [37,38] can implement the statistical significance of the coincident events. Originally, the method is based on the signal-to-noise ratio (SNR) of the matched filter, but in our case, we could incarnate it with the correlation score instead of using the SNR. Then we could conduct the timeslide by shifting a stride back and forth for a given high-scored event, counting the number of events.

Qualitatively, an alternative way is to refer to the subsystems that affect each other. When we investigate a family of correlated subsystems that is closely related, because some similar observations can be seen from the different related channels, one can estimate the observed correlation with a reliable statistical significance, at least, qualitatively. More precisely, when we investigate the effect of magnetic fields, the magnetometer-related channels with a family of the same subsystem (MAG-subsystem) are affected with similar behaviors, i.e., a similar correlation appears. Therefore, we could conclude that the correlation index in the magnetometer-related channels reveals a piece
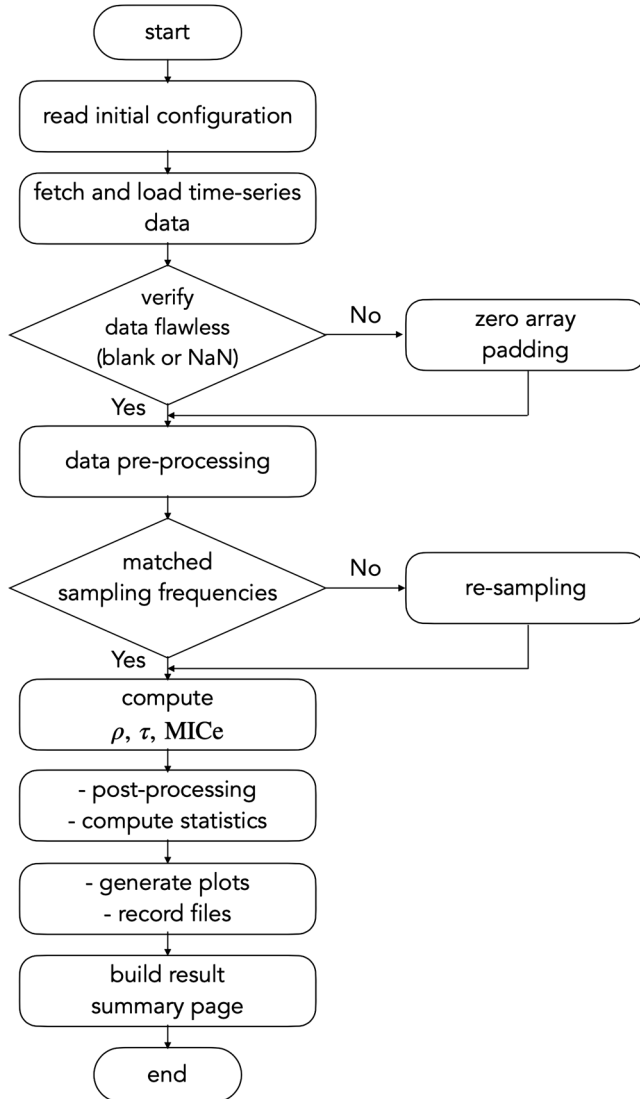
TABLE I. The channel information and parameters of MICe used in this study: these values are selected by the empirical optimized parameter selection in [36].

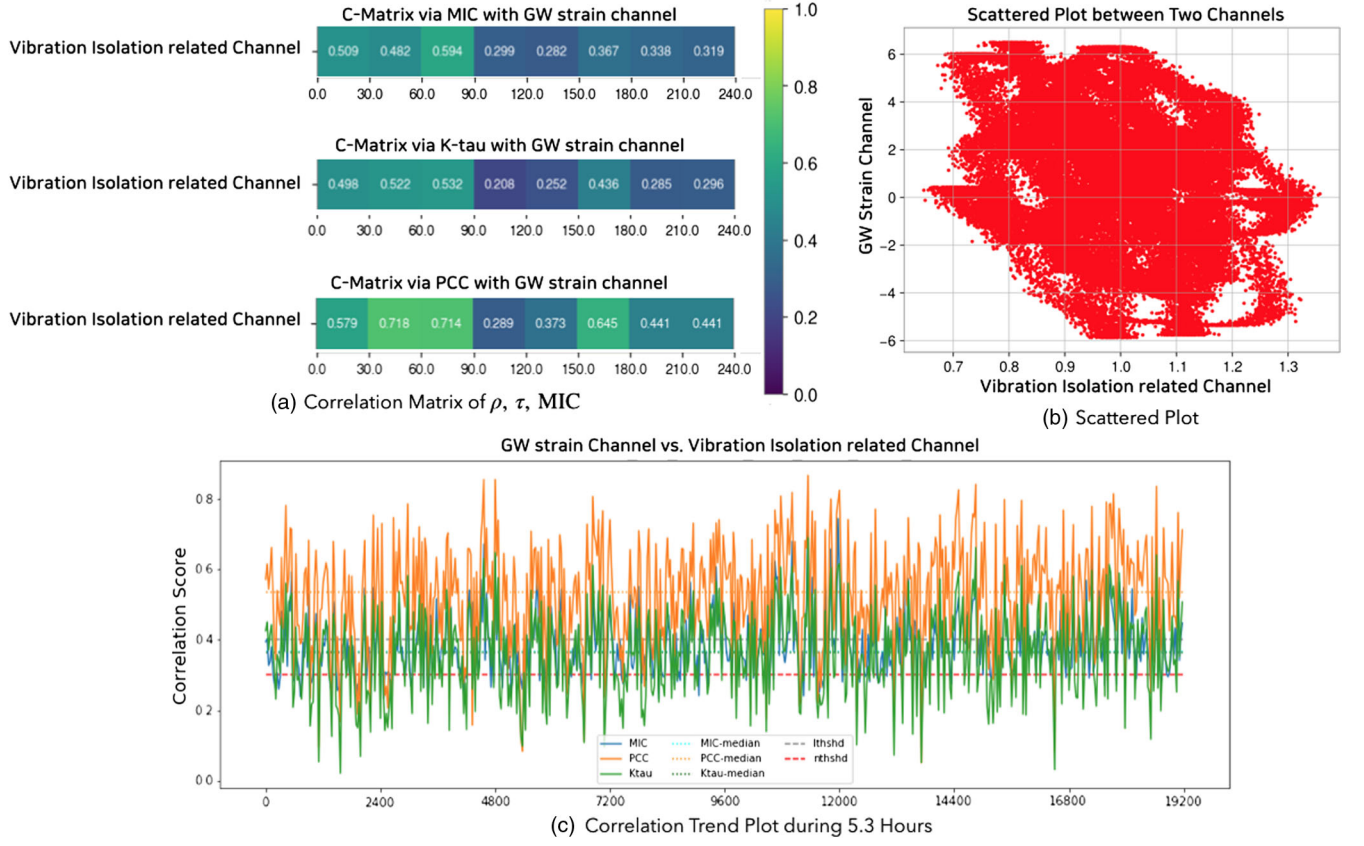| Channel type	Stride(s)	Sampling rate (Hz)	Data size	α	c
GW strain	2	4096	8192	0.5	0.7
BNS range	512	16	8192	0.5	7.0

FIG. 2. Exemplary plots provided by the *CAGMon* tool: (a) correlation matrix plot of ρ, τ, and MIC between GW strain channel and auxiliary channels, (b) scattered plot of both channels (each relative amplitude is rescaled by its median), and (c) correlation trend plot within a certain duration. Here, we demonstrate one of the vibration isolation system (VIS) channels among tens of auxiliary channels in the KAGRA GW detector as a simple example.

of reliable evidence with the statistical significance of the observation. A specific example can be found in the applications in the following section.

III. APPLICATION TO GRAVITATIONAL-WAVE DATA

The *CAGMon* tool was designed to identify and diagnose associated auxiliary channels that influence the GW strain channel for GW detection, producing a time-sequential correlation trend propagated by instrumental or environmental disturbances. We can estimate a significant association as excess from the long duration trend of its median value. In this section, we apply the *CAGMon* tool to the KAGRA GW data for characterizing and identifying associations between channels caused by well-known environmental and instrumental events such as lightning stroke and air compressor noises.

A. Magnetic field transients from lightning stroke

First, the *CAGMon* tool can be applied to investigate a magnetic field transient noise caused by a lightning stroke event observed by the KAGRA collaboration that was verified by a clear mutual relationship between two
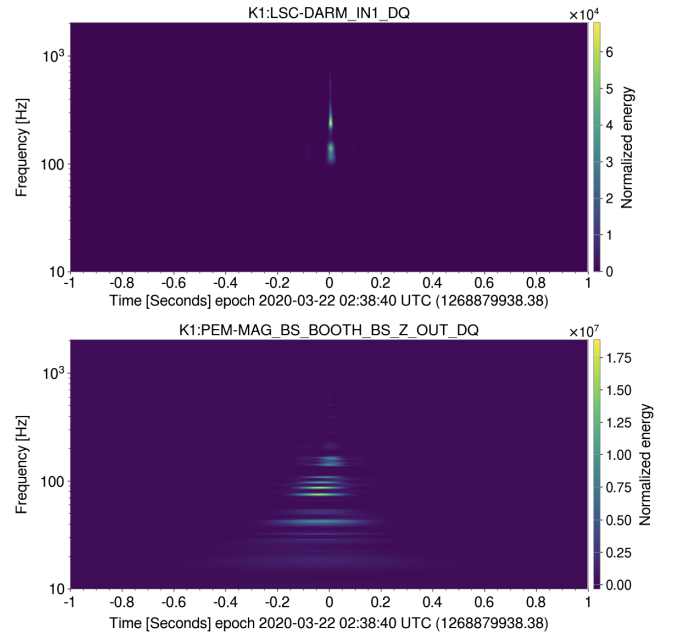

FIG. 3. Omega scan plots of the GW strain channel (top) and the magnetometer channel at the beam splitter (BS) station (bottom).

channels in [39]. A lightning stroke is a locally enormous release of electromagnetic waves, which clearly affects the GW detector by a sudden variance of magnetic fields. Note that the relevant studies in LIGO, Virgo, and KAGRA are shown in terms of a short duration magnetic fields effect [38], the amplified magnetic effect in the underground site [40], and the transient magnetic effects caused by a large scale lightning strike during GW150914 [41].

The KAGRA GW detector was influenced by a lightning stroke at 02:38:40.38 on March 21, 2020, UTC, which was recorded in the GW strain channel and magnetometers. This detection was the first evidence observed by the KAGRA GW detector that the lightning strokes in the atmosphere would be able to affect the underground-based GW detector within the detection range. The disturbance caused by the electromagnetic fields from the lightning stroke can affect the GW strain channel of the KAGRA GW detector, producing a coincident transient noise in the strain channel with a significant association. Hence, we can infer that there exists a significant correlation between the GW strain channel and the magnetometer-related channels. The omega-scan spectrogram plots based on Q-transform are presented in Fig. 3.

Here, we used two seconds stride of data with 8192 data size from the GW strain channel to investigate a correlation between short-duration data segments. The mutual correlations between the GW strain channel and the magnetometer-related channels detect a meaningful signal at the event time of the lightning stroke. Here, correlation trend plots are depicted in Fig. 4, yielding a clear peak signal at the event time of the lightning stroke. It is inferred that the magnetic field noise from the lightning stroke event can affect the GW strain channel. Interestingly,
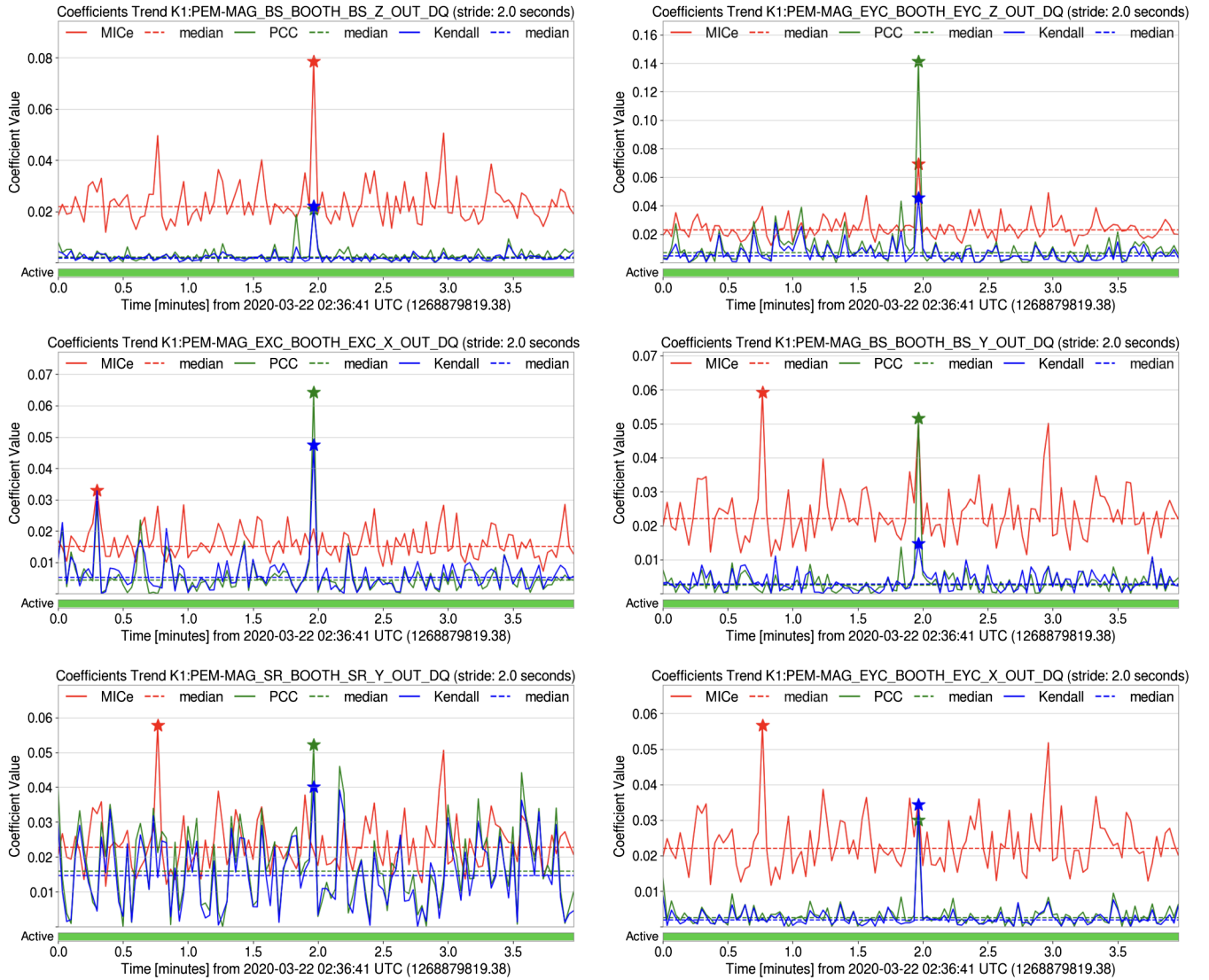

FIG. 4.   Correlation trend plot of association between the GW interferometer channel and magnetometer associated channels at the BS station at the lightning event time. The solid line and the dashed line indicate the strength of the correlation coefficient and median value, respectively. Here, the star mark represents the maximum value of $(\mathrm{MICe}, \rho, \tau)$.

TABLE II. Correlation values of MICe, ρ, and τ at the lightning stroke event time and some associated channels with the magnetometer in the KAGRA GW detector.

Event time (GPS)	Associated auxiliary channels[a]	MICe	Med(MICe)[b] (10^{-2})	ρ	Med(ρ) (10^{-2})	τ	Med(τ) (10^{-2})
	K1:PEM[c]-MAG_BS_BOOTH_BS_Z_OUT_DQ	0.079	2.210	0.021	0.212	0.022	0.172
March 22, 2020	K1:PEM-MAG_BS_BOOTH_BS_Y_OUT_DQ	0.050	2.210	0.052	0.250	0.015	0.275
02:38:39-41UTC	K1:PEM-MAG_BS_BOOTH_BS_X_OUT_DQ	0.026	2.188	0.040	0.266	0.001	0.197
(1268879937.38	K1:PEM-MAG_EXC_BOOTH_EXC_X_OUT_DQ	0.021	1.499	0.064	0.423	0.047	0.521
−1268879939.38)	K1:PEM-MAG_EYC_BOOTH_EYC_Z_OUT_DQ	0.069	2.309	0.141	0.709	0.045	0.474
	K1:PEM-MAG_SR_BOOTH_SR_Z_OUT_DQ	0.022	2.271	0.052	1.595	0.040	1.458

[a]All exhibited channels here are related to magnetometers installed in several areas nearby the KAGRA GW detector.
[b]The median value was computed for four minutes duration segment.
[c]PEM stands for the physical environment monitor that is related to subsystems for sensing environmental changes such as seismomter, magnetometer, accelerometer, and so on [42].

the aspect of each correlation exhibited in Fig. 4 shows that they are linearly or nonlinearly correlated with each other for the individual subsystems. The values with significant correlations are much greater than the median values of four minutes duration data segment. In addition, we list the magnetometer-associated subsystems with a significant correlation to the lightning stroke event in Table. II.

B. Periodic noises from air compressors

We applied the *CAGMon* tool to the observing run of KAGRA (O3 GK run) from April 7, 2020, to April 21, 2020 [43,44]. Unlike the aforementioned analysis for transient noise, we used the binary neutron star (BNS) range channel instead of the GW strain channel to investigate the coherence between the sudden drops in the BNS range plot and the air compressor noises. The BNS range channel is a monitoring channel that represents the volume/ orientation-averaged distance based on the GW signals with a matched filtered signal-to-noise ratio (SNR) of 8 from the two 1.4 $M_\odot$ neutron stars in a GW detector [5,45]. The slowly moving behavior caused by long-duration noise disturbances in the LIGO detector has been investigated via the least absolute shrinkage and selection operator regression, where the BNS range channel was used [24].

In this analysis, we examined internal or external influences on the detector's sensitivity during the long period of observational mode. Because data record a minute trend of sensitivity range with around three minutes latency, we set 512 seconds of stride and 8192 of data size. With this configuration, significant correlations were periodically repeated between the BNS range channel and microphone channels near the beam splitter (BS) station. These correlated peak signals have a harmonic frequency of 26.5 Hz and repeat with a periodicity of 2.58 hours every day. Figure 5 exhibits the correlation trend plots of this periodic behavior using the *CAGMon* tool. In addition, it has been found that the air compressor at the output mode cleaner (OMC) produces periodic vibrations that correlated through nearby instruments to result in the sudden drop of

the BNS range of the KAGRA sensitivity curve. The BNS range drop owing to the periodical air compressor noises is demonstrated in Fig. 6. This phenomenon was previously reported during the KAGRA engineering operation and reduced significantly via the installation of vibration isolation and soundproof devices in the air compressors. However, the *CAGMon* tool caught this effect even if it had been reduced.

Consequently, this example indicates that association may exist if the trend of the coefficients is changed regardless of their strength.
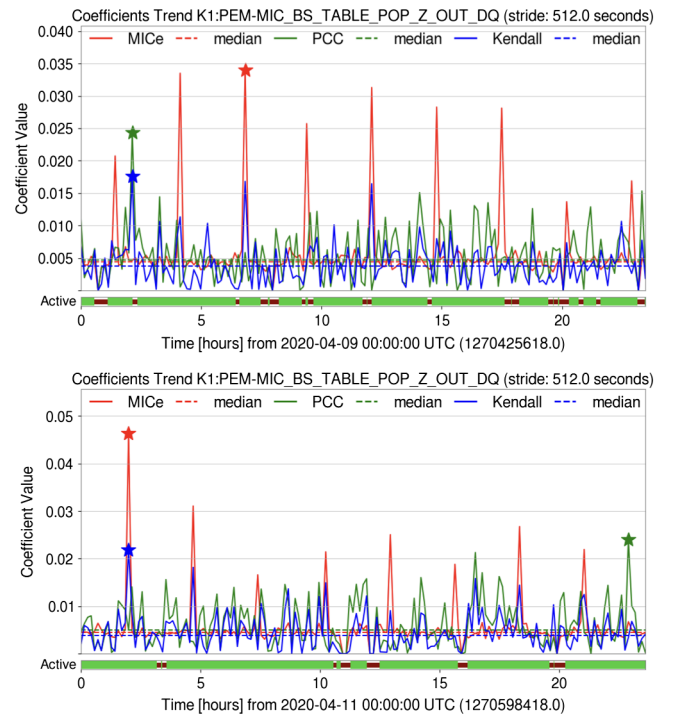

FIG. 5.   Correlation trend plots between the BNS range and microphone channels installed in the BS station Pick-off Port room on April 9, 2020 (top) and April 11, 2020 (bottom); relatively strong MICe correlations were repeated every 2.58 hours.
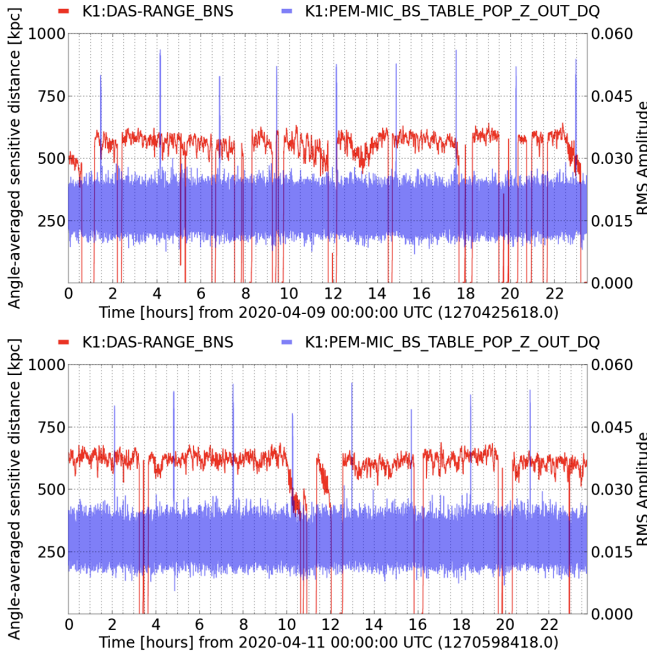
FIG. 6. The BNS range curve experiences a sudden drop of as much as about 50 kpc owing to the air compressor vibration from the OMC station on April 9, 2020 (top) and April 11, 2020 (bottom).

C. Glimpse of acoustic noise from winds

Another interesting point via the *CAGMon* tool was observed by a clear correlation between the BNS range channel and the microphone/vibration isolation channels. The correlation trend plot exhibits a meaningful association from morning (9 a.m., JST) to evening (7 p.m., JST) every day in the spring season, which is inferred by a relationship with strong wind power during this time (Fig. 7). Because the KAGRA GW detector is installed inside the underground tunnel in Mt. Ikeno, we infer that this is responsible for the propagation of an acoustic wave noise owing to strong winds between Mt. Ikeno. Figure 7 exhibits an interesting feature of MICe because this association only was detected by MIC rather than PCC and Ktau; one can infer that the effect of winds can affect nonlinearly the KAGRA detector. A theoretical possibility of influences by the fluctuating gravity gradient noises in the ground and air around the detector has been studied in [46]. For this reason, it is worthwhile studying because the seismic coupling and up-conversion effect may affect the GW detection band for the underground detector.

A potential scenario of this effect is considered by a simple simulation of the elastic wave equation in the underground like the KAGRA detector. The wall of the L-shape tunnel in the KAGRA detector consists of the concrete material and air inside the tunnel. The *y* arm
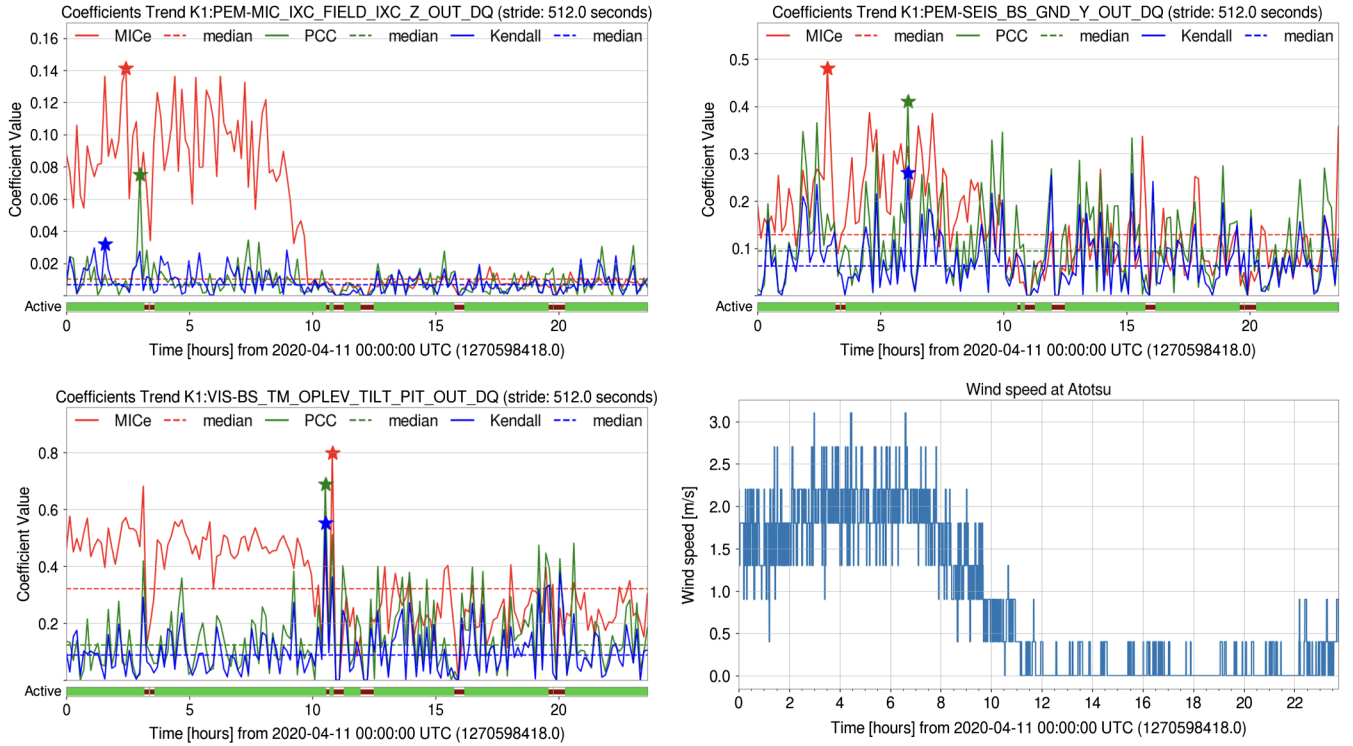

FIG. 7. Correlation trend plots between the BNS range channel and microphone channel installed in the initial x-arm tank (top left), the test mass optical leverage at the BS VIS channel (bottom left), and the y-axis seismometer at BS (top right). The plot of wind speed is defined as the average wind speed over two minutes at the entrance of KAGRA (Atotsu) on April 11, 2020 (bottom right).
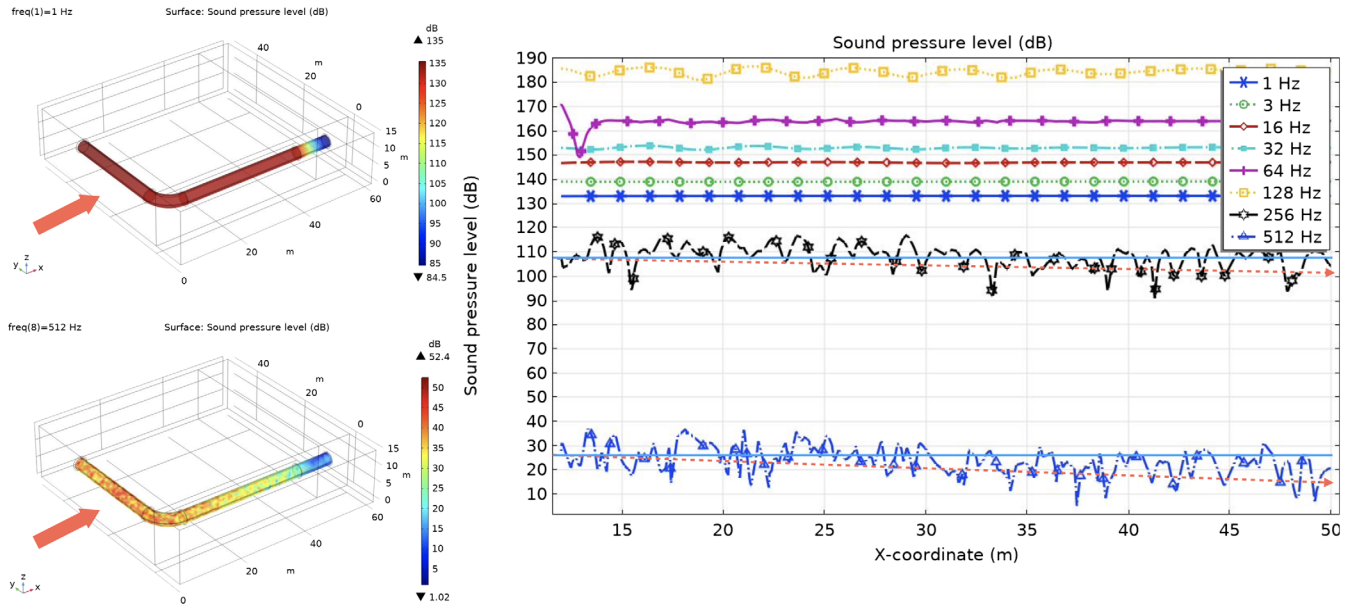
FIG. 8. The simulation of seismic and acoustic wave propagation in the underground tunnel of KAGRA and its x-arm sound pressure level in diverse frequencies. The perfectly matched layer was applied in the end region of the x arm.

of the KAGRA detector is parallel to the slop of the valley in Mt. Ikeno whereas the x arm is located from the slope to deep inside the mountain. The seismic incident wave generated from the strong winds striking the slope vibrates and propagates then transforms into an acoustic wave inside the tunnel. This acoustic vibration produces sound pressure noise inside the tunnel. Because of the location of both arms in the KAGRA detector, the sound pressure noise level in the y-arm direction is much more severe than that in the x-arm direction due to the attenuation inside the deeper region of the tunnel.

The finite element method simulation of the seismic and acoustic waves with a multiphysics configuration and the sound pressure level along the x-arm direction in diverse frequencies are shown in Fig. 8.[1] Consequently, the seismic vibration from the slope can propagate to air fluctuation in the tunnel of the KAGRA, then producing acoustic noises inside the tunnel, which needs to be verified by more accurate simulation and instrumental measurement.

IV. DISCUSSIONS

We developed a novel tool for identifying and diagnosing data association between two variables to investigate presumably correlated events between multichannels of the GW detectors, called *CAGMon*. In this tool, three linear and nonlinear measures, PCC, Ktau, and MIC, were adopted and the optimized parameter selection for MIC was referred to in [36].

We applied this tool for the transient and periodic noises caused by a lightning stroke event and an acoustic noise caused by an air compressor in the KAGRA GW detector, respectively. Consequently, we verified that several magnetometer-associated subsystems influenced by the lightning stroke event have a significant association with the GW strain channel. On the other hand, we found that the sudden detection range reductions in the BNS range curve of the KAGRA detector are associated with the periodic noise every 2.58 hours appearing in the microphone channels. The noise was identified with a harmonic frequency of 26.5 Hz and the evident cause was confirmed by acoustic noises from the air compressor nearby the BS station. Finally, it observed a glimpse of the wind effect on the underground detector via the *CAGMon* tool. In the daytime, the air in nearby mountains becomes heated and makes a relatively strong wind between mountains. This wind hits the ground surface where the detector is located, yielding and propagating microseismic noise and infrasound waves toward the underground detector [47]. Thus, the detector experiences acoustic and seismic vibrations during windy times. This scenario seems plausible but needs to verify in various aspects. Consequently, the seismic vibration from the slope can propagate to air fluctuation in the tunnel of the KAGRA, then producing acoustic noises inside the tunnel, which needs to be verified by more accurate simulation and instrumental measurement.

Potentially, the *CAGMon* tool and its application will help to overcome several limitations in the KAGRA detector, and thereby contribute not only to improving the noise reduction study but also to developing advanced tools and interfaces for the next-generation gravitational-wave telescopes.

---

[1]The simulation has been performed by COMSOL-Multiphysics 5.6 with somewhat *ad hoc* parameters and assumptions to check the possibility of the scenario.

Furthermore, this tool will help scientists in the GW detection as well as other fields of science.

ACKNOWLEDGMENTS

[1] B. P. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), Phys. Rev. Lett. **116**, 061102 (2016).

[2] J. Aasi *et al.* (LIGO Scientific Collaboration), Classical Quantum Gravity **32**, 074001 (2015).

[3] F. Acernese *et al.* (Virgo Collaboration), Classical Quantum Gravity **32**, 024001 (2015).

[4] T. Akutsu *et al.* (KAGRA Collaboration), Nat. Astron. **3**, 35 (2019).

[5] B. P. Abbott *et al.* (KAGRA, LIGO Scientific, Virgo Collaborations), Living Rev. Relativity **21**, 3 (2018).

[6] M. Punturo *et al.*, Classical Quantum Gravity **27**, 194002 (2010).

[7] B. P. Abbott *et al.* (LIGO Scientific Collaboration), Classical Quantum Gravity **34**, 044001 (2017).

[8] J. Aasi *et al.* (LIGO Scientific, Virgo Collaborations), Classical Quantum Gravity **32**, 115012 (2015).

[9] B. P. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), Classical Quantum Gravity **33**, 134001 (2016).

[10] P.-J. Jung, K.-Y. Kim, J. J. Oh, S. H. Oh, E. J. Son, and Y.-M. Kim, J. Korean Phys. Soc. **73**, 1197 (2018).

[11] P. B. Covas *et al.* (LSC Collaboration), Phys. Rev. D **97**, 082002 (2018).

[12] D. Davis *et al.* (LIGO Collaboration), Classical Quantum Gravity **38**, 135014 (2021).

[13] N. Christensen (LIGO Scientific, Virgo Collaborations), Classical Quantum Gravity **27**, 194010 (2010).

[14] J. Aasi *et al.* (LIGO Scientific, Virgo Collaborations), Classical Quantum Gravity **32**, 115012 (2015).

[15] J. Aasi *et al.* (Virgo Collaboration), Classical Quantum Gravity **29**, 155002 (2012).

[16] F. Acernese *et al.*, Classical Quantum Gravity **24**, S433 (2007).

[17] L. Nuttall *et al.*, Classical Quantum Gravity **32**, 245005 (2015).

[18] T. Isogai (LIGO Scientific, Virgo Collaborations), J. Phys. Conf. Ser. **243**, 012005 (2010).

[19] T. Prestegard, E. Thrane, N. L. Christensen, M. W. Coughlin, B. Hubbert, S. Kandhasamy, E. MacAyeal, and V. Mandic, Classical Quantum Gravity **29**, 095018 (2012).

[20] E. Thrane, S. Kandhasamy, C. D. Ott, W. G. Anderson, N. L. Christensen, M. W. Coughlin, S. Dorsher, S. Giampanis, V. Mandic, A. Mytidis, T. Prestegard, P. Raffai, and B. Whiting, Phys. Rev. D **83**, 083004 (2011).

[21] F. Robinet, N. Arnaud, N. Leroy, A. Lundgren, D. Macleod, and J. McIver, SoftwareX **12**, 100620 (2020).

[22] E. J. Son, W. Kim, Y.-M. Kim, J. McIver, J. J. Oh, and S. H. Oh, J. Korean Phys. Soc. **78**, 878 (2021).

[23] P. Jaranowski and A. Krolak, Living Rev. Relativity **8**, 3 (2005).

[24] M. Walker, A. F. Agnew, J. Bidler, A. Lundgren, A. Macedo, D. Macleod, T. J. Massinger, O. Patane, and J. R. Smith, Classical Quantum Gravity **35**, 225002 (2018).

[25] E. Cuoco *et al.*, Mach. Learn. Sci. Tech. **2**, 011002 (2021).

[26] S. Bose, B. Hall, N. Mazumder, S. Dhurandhar, A. Gupta, and A. Lundgren, J. Phys. Conf. Ser. **716**, 012007 (2016).

[27] P. Ajith, T. Isogai, N. Christensen, R. X. Adhikari, A. B. Pearlman, A. Wein, A. J. Weinstein, and B. Yuan, Phys. Rev. D **89**, 122001 (2014).

[28] K. Pearson, Biometrika **13**, 25 (1920).

[29] M. G. Kendall, Biometrika **30**, 81 (1938).

[30] D. N. Reshef, Y. A. Reshef, H. K. Finucane, S. R. Grossman, G. McVean, P. J. Turnbaugh, E. S. Lander, M. Mitzenmacher, and P. C. Sabeti, Science **334**, 1518 (2011).

[31] D. N. Reshef, Y. A. Reshef, P. C. Sabeti, and M. Mitzenmacher, Ann. Appl. Stat. **12**, 123 (2018).

[32] D. N. Reshef, Y. A. Reshef, P. C. Sabeti, and M. Mitzenmacher, Ann. Appl. Stat. **12**, 123 (2018).

[33] T. Akutsu *et al.*, Prog. Theor. Exp. Phys. **2021**, 05A101 (2020).

[34] K. Kokeyama (KAGRA Collaboration), in *3rd World Summit on Exploring the Dark Side of the Universe* (University of Kansas Libraries, Lawrence, 2020).

[35] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley Series in Telecommunications and Signal Processing, 2nd ed. (Wiley-Interscience, New York, 2006).

[36] P. Jung, S. H. Oh, E. J. Son, Y.-M. Kim, and J. J. Oh, Prog. Theor. Exp. Phys. **2022**, 073H01 (2022).

[37] M. Was, M.-A. Bizouard, V. Brisson, F. Cavalier, M. Davier, P. Hello, N. Leroy, F. Robinet, and M. Vavoulidis, Classical Quantum Gravity **27**, 015005 (2010).

[38] I. Kowalska-Leszczynska *et al.*, Classical Quantum Gravity **34**, 074002 (2017).

[39] T. Washimi, T. Yokozawa, M. Nakano, T. Tanaka, K. Kaihotsu, Y. Mori, and T. Narita, J. Instrum. **16**, P07033 (2021).

[40] S. Atsuta *et al.*, J. Phys. Conf. Ser. **716**, 012020 (2016).

[41] B. P. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), Classical Quantum Gravity **33**, 134001 (2016).

[42] T. Akutsu *et al.*, Prog. Theor. Exp. Phys. **2021**, 05A102 (2021).

[43] R. Abbott *et al.* (LIGO Scientific, Virgo, KAGRA Collaborations), Prog. Theor. Exp. Phys. **2022**, 063F01 (2022).

[44] H. Abe *et al.* (KAGRA Collaboration), arXiv:2203.07011.

[45] L. S. Finn and D. F. Chernoff, Phys. Rev. D **47**, 2198 (1993).

[46] T. Creighton, Classical Quantum Gravity **25**, 125011 (2008).

[47] G. Chimonas, J. Atmos. Sci. **34**, 806 (1977).