PTEP

Prog. Theor. Exp. Phys. **2020**, 053F01 (19 pages)
DOI: 10.1093/ptep/ptaa056

Application of independent component analysis to the iKAGRA data

KAGRA Collaboration

T. Akutsu[1,2], M. Ando[3,4,1], K. Arai[5], Y. Arai[5], S. Araki[6], A. Araya[7], N. Aritomi[3], H. Asada[8], Y. Aso[9,10], S. Atsuta[11], K. Awai[12], S. Bae[13], Y. Bae[14], L. Baiotti[15], R. Bajpai[16], M. A. Barton[1], K. Cannon[4], E. Capocasa[1], M. Chan[17], C. Chen[18,19], K. Chen[20], Y. Chen[19], H. Chu[20], Y-K. Chu[21], K. Craig[5], W. Creus[21], K. Doi[22], K. Eda[4], S. Eguchi[17], Y. Enomoto[3], R. Flaminio[23,1], Y. Fujii[24], M.-K. Fujimoto[1], M. Fukunaga[5], M. Fukushima[1], T. Furuhata[22], G. Ge[25], A. Hagiwara[5,26], S. Haino[21], K. Hasegawa[5], K. Hashino[22], H. Hayakawa[12], K. Hayama[17], Y. Himemoto[27], Y. Hiranuma[28], N. Hirata[1], S. Hirobayashi[29], E. Hirose[5], Z. Hong[30], B. H. Hsieh[31], G-Z. Huang[30], P. Huang[25], Y. Huang[21], B. Ikenoue[1], S. Imam[30], K. Inayoshi[32], Y. Inoue[20], K. Ioka[33], Y. Itoh[34,35], K. Izumi[36], K. Jung[37], P. Jung[12], T. Kaji[35], T. Kajita[38], M. Kakizaki[22], M. Kamiizumi[12], S. Kanbara[22], N. Kanda[34,35], S. Kanemura[39], M. Kaneyama[35], G. Kang[13], J. Kasuya[11], Y. Kataoka[11], K. Kawaguchi[5], N. Kawai[11], S. Kawamura[12], T. Kawasaki[3], C. Kim[40], J. C. Kim[41], W. S. Kim[14], Y.-M. Kim[37], N. Kimura[26], T. Kinugawa[42], S. Kirii[12], N. Kita[3], Y. Kitaoka[35], H. Kitazawa[22], Y. Kojima[43], K. Kokeyama[12], K. Komori[3], A. K. H. Kong[19], K. Kotake[17], C. Kozakai[9], R. Kozu[44], R. Kumar[45], J. Kume[4,3,*], C. Kuo[20], H-S. Kuo[30], S. Kuroyanagi[46], K. Kusayanagi[11], K. Kwak[37], H. K. Lee[47], H. M. Lee[48,49], H. W. Lee[41], R. Lee[19], M. Leonardi[1], C. Lin[37], C-Y. Lin[50], F-L. Lin[30], G. C. Liu[18], Y. Liu[51], L. Luo[21], E. Majorana[52], S. Mano[53], M. Marchio[1], T. Matsui[54], F. Matsushima[22], Y. Michimura[3], N. Mio[55], O. Miyakawa[12], A. Miyamoto[35], T. Miyamoto[44], Y. Miyazaki[3], K. Miyo[12], S. Miyoki[12], W. Morii[56], S. Morisaki[4], Y. Moriwaki[22], T. Morozumi[5], M. Musha[57], K. Nagano[5], S. Nagano[58], K. Nakamura[1], T. Nakamura[59], H. Nakano[60], M. Nakano[22,5], K. Nakao[35], R. Nakashima[11], T. Narikawa[59], L. Naticchioni[52], R. Negishi[28], L. Nguyen Quynh[61], W.-T. Ni[25,62,63], A. Nishizawa[4], Y. Obuchi[1], T. Ochi[5], W. Ogaki[5], J. J. Oh[14], S. H. Oh[14], M. Ohashi[12], N. Ohishi[9], M. Ohkawa[64], K. Okutomi[12], K. Oohara[28], C. P. Ooi[3], S. Oshino[12], K. Pan[19], H. Pang[20], J. Park[65], F. E. Peña Arellano[12], I. Pinto[66], N. Sago[67], M. Saijo[68], S. Saito[1], Y. Saito[12], K. Sakai[69], Y. Sakai[28], Y. Sakai[3], Y. Sakuno[17], M. Sasaki[70], Y. Sasaki[71], S. Sato[72], T. Sato[64], T. Sawada[34], T. Sekiguchi[4], Y. Sekiguchi[73], N. Seto[59], S. Shibagaki[17], M. Shibata[33,74], R. Shimizu[1], T. Shimoda[3], K. Shimode[12], H. Shinkai[75], T. Shishido[10], A. Shoda[1], K. Somiya[11], E. J. Son[14], H. Sotani[1], A. Suemasa[57], R. Sugimoto[22], T. Suzuki[64], T. Suzuki[5], H. Tagoshi[5], H. Takahashi[71], R. Takahashi[1], A. Takamori[7], S. Takano[3], H. Takeda[3], M. Takeda[28], H. Tanaka[31], K. Tanaka[35], K. Tanaka[31], T. Tanaka[5], T. Tanaka[59], S. Tanioka[1,10], E. N. Tapia San Martin[1], D. Tatsumi[1], S. Telada[76], T. Tomaru[1], Y. Tomigami[34], T. Tomura[12], F. Travasso[77,78], L. Trozzo[12], T. Tsang[79], K. Tsubono[3], S. Tsuchida[34], T. Tsuzuki[1], D. Tuyenbayev[21], N. Uchikata[80], T. Uchiyama[12], A. Ueda[26], T. Uehara[81,82], S. Ueki[71], K. Ueno[4], G. Ueshima[71], F. Uraguchi[1], T. Ushiba[5], M. H. P. M. van Putten[83], H. Vocca[78], S. Wada[3], T. Wakamatsu[28], J. Wang[25], C. Wu[19], H. Wu[19], S. Wu[19], W-R. Xu[30], T. Yamada[31], A. Yamamoto[6], K. Yamamoto[22], K. Yamamoto[31], S. Yamamoto[75], T. Yamamoto[12], K. Yokogawa[22], J. Yokoyama[4,3], T. Yokozawa[12], T. H. Yoon[84], T. Yoshioka[22], H. Yuzurihara[5], S. Zeidler[1], Y. Zhao[1], and Z.-H. Zhu[85]

[1]*Gravitational Wave Project Office, National Astronomical Observatory of Japan (NAOJ), Mitaka City, Tokyo 181-8588, Japan*
[2]*Advanced Technology Center, National Astronomical Observatory of Japan (NAOJ), Mitaka City, Tokyo 181-8588, Japan*
[3]*Department of Physics, The University of Tokyo, Bunkyo-ku, Tokyo 113-0033, Japan*
[4]*Research Center for the Early Universe (RESCEU), The University of Tokyo, Bunkyo-ku, Tokyo 113-0033, Japan*
[5]*Institute for Cosmic Ray Research (ICRR), KAGRA Observatory, The University of Tokyo, Kashiwa City, Chiba 277-8582, Japan*
[6]*Accelerator Laboratory, High Energy Accelerator Research Organization (KEK), Tsukuba City, Ibaraki 305-0801, Japan*
[7]*Earthquake Research Institute, The University of Tokyo, Bunkyo-ku, Tokyo 113-0032, Japan*
[8]*Department of Mathematics and Physics, Hirosaki University, Hirosaki City, Aomori 036-8561, Japan*
[9]*Kamioka Branch, National Astronomical Observatory of Japan (NAOJ), Kamioka-cho, Hida City, Gifu 506-1205, Japan*
[10]*The Graduate University for Advanced Studies (SOKENDAI), Mitaka City, Tokyo 181-8588, Japan*
[11]*Graduate School of Science and Technology, Tokyo Institute of Technology, Meguro-ku, Tokyo 152-8551, Japan*
[12]*Institute for Cosmic Ray Research (ICRR), KAGRA Observatory, The University of Tokyo, Kamioka-cho, Hida City, Gifu 506-1205, Japan*
[13]*Korea Institute of Science and Technology Information (KISTI), Yuseong-gu, Daejeon 34141, Korea*
[14]*National Institute for Mathematical Sciences, Daejeon 34047, Korea*
[15]*Department of Earth and Space Science, Graduate School of Science, Osaka University, Toyonaka City, Osaka 560-0043, Japan*
[16]*School of High Energy Accelerator Science, The Graduate University for Advanced Studies (SOKENDAI), Tsukuba City, Ibaraki 305-0801, Japan*
[17]*Department of Applied Physics, Fukuoka University, Jonan, Fukuoka City, Fukuoka 814-0180, Japan*
[18]*Department of Physics, Tamkang University, Danshui Dist., New Taipei City 25137, Taiwan*
[19]*Department of Physics and Institute of Astronomy, National Tsing Hua University, Hsinchu 30013, Taiwan*
[20]*Department of Physics, Center for High Energy and High Field Physics, National Central University, Zhongli District, Taoyuan City 32001, Taiwan*
[21]*Institute of Physics, Academia Sinica, Nankang, Taipei 11529, Taiwan*
[22]*Department of Physics, University of Toyama, Toyama City, Toyama 930-8555, Japan*
[23]*Univ. Grenoble Alpes, Laboratoire d'Annecy de Physique des Particules (LAPP), Université Savoie Mont Blanc, CNRS/IN2P3, F-74941 Annecy, France*
[24]*Department of Astronomy, The University of Tokyo, Mitaka City, Tokyo 181-8588, Japan*
[25]*State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, Wuhan Institute of Physics and Mathematics (WIPM), Chinese Academy of Sciences, Xiaohongshan, Wuhan 430071, China*
[26]*Applied Research Laboratory, High Energy Accelerator Research Organization (KEK), Tsukuba City, Ibaraki 305-0801, Japan*
[27]*College of Industrial Technology, Nihon University, Narashino City, Chiba 275-8575, Japan*
[28]*Graduate School of Science and Technology, Niigata University, Nishi-ku, Niigata City, Niigata 950-2181, Japan*
[29]*Faculty of Engineering, University of Toyama, Toyama City, Toyama 930-8555, Japan*
[30]*Department of Physics, National Taiwan Normal University, sec. 4, Taipei 116, Taiwan*
[31]*Institute for Cosmic Ray Research (ICRR), Research Center for Cosmic Neutrinos (RCCN), The University of Tokyo, Kashiwa City, Chiba 277-8582, Japan*
[32]*Kavli Institute for Astronomy and Astrophysics, Peking University, Yiheyuan Road 5, Haidian District, Beijing 100871, China*
[33]*Yukawa Institute for Theoretical Physics (YITP), Kyoto University, Sakyou-ku, Kyoto City, Kyoto 606-8502, Japan*
[34]*Department of Physics, Graduate School of Science, Osaka City University, Sumiyoshi-ku, Osaka City, Osaka 558-8585, Japan*
[35]*Nambu Yoichiro Institute of Theoretical and Experimental Physics (NITEP), Osaka City University, Sumiyoshi-ku, Osaka City, Osaka 558-8585, Japan*
[36]*Institute of Space and Astronautical Science (JAXA), Chuo-ku, Sagamihara City, Kanagawa 252-0222, Japan*

[37]*Department of Physics, School of Natural Science, Ulsan National Institute of Science and Technology (UNIST), Ulsan 44919, Korea*
[38]*Institute for Cosmic Ray Research (ICRR), The University of Tokyo, Kashiwa City, Chiba 277-8582, Japan*
[39]*Graduate School of Science, Osaka University, Toyonaka City, Osaka 560-0043, Japan*
[40]*Department of Physics, Ewha Womans University, Seodaemun-gu, Seoul 03760, Korea*
[41]*Department of Computer Simulation, Inje University, Gimhae, Gyeongsangnam-do 50834, Korea*
[42]*Department of Astronomy, The University of Tokyo, Bunkyo-ku, Tokyo 113-0033, Japan*
[43]*Department of Physical Science, Hiroshima University, Higashihiroshima City, Hiroshima 903-0213, Japan*
[44]*Institute for Cosmic Ray Research (ICRR), Research Center for Cosmic Neutrinos (RCCN), The University of Tokyo, Kamioka-cho, Hida City, Gifu 506-1205, Japan*
[45]*California Institute of Technology, Pasadena, CA 91125, USA*
[46]*Institute for Advanced Research, Nagoya University, Furocho, Chikusa-ku, Nagoya City, Aichi 464-8602, Japan*
[47]*Department of Physics, Hanyang University, Seoul 133-791, Korea*
[48]*Korea Astronomy and Space Science Institute (KASI), Yuseong-gu, Daejeon 34055, Korea*
[49]*Department of Physics and Astronomy, Seoul National University, Gwanak-gu, Seoul 08826, Korea*
[50]*National Center for High-performance Computing, National Applied Research Laboratories, Hsinchu Science Park, Hsinchu City 30076, Taiwan*
[51]*Department of Advanced Materials Science, The University of Tokyo, Kashiwa City, Chiba 277-8582, Japan*
[52]*Istituto Nazionale di Fisica Nucleare (INFN), Sapienza University, Roma 00185, Italy*
[53]*Department of Mathematical Analysis and Statistical Inference, The Institute of Statistical Mathematics, Tachikawa City, Tokyo 190-8562, Japan*
[54]*School of Physics, Korea Institute for Advanced Study (KIAS), Seoul 02455, Korea*
[55]*Institute for Photon Science and Technology, The University of Tokyo, Bunkyo-ku, Tokyo 113-8656, Japan*
[56]*Disaster Prevention Research Institute, Kyoto University, Uji City, Kyoto 611-0011, Japan*
[57]*Institute for Laser Science, University of Electro-Communications, Chofu City, Tokyo 182-8585, Japan*
[58]*The Applied Electromagnetic Research Institute, National Institute of Information and Communications Technology (NICT), Koganei City, Tokyo 184-8795, Japan*
[59]*Department of Physics, Kyoto University, Sakyou-ku, Kyoto City, Kyoto 606-8502, Japan*
[60]*Faculty of Law, Ryukoku University, Fushimi-ku, Kyoto City, Kyoto 612-8577, Japan*
[61]*Department of Physics, University of Notre Dame, Notre Dame, IN 46556, USA*
[62]*Department of Physics, National Tsing Hua University, Hsinchu 30013, Taiwan*
[63]*School of Optical Electrical and Computer Engineering, The University of Shanghai for Science and Technology, Shanghai 200093, China*
[64]*Faculty of Engineering, Niigata University, Nishi-ku, Niigata City, Niigata 950-2181, Japan*
[65]*Optical instrument developement team, Korea Basic Science Institute, 169-148, Gwahak-ro, Yuseong-gu, Daejeon, Korea*
[66]*Department of Engineering, University of Sannio, Benevento 82100, Italy*
[67]*Faculty of Arts and Science, Kyushu University, Nishi-ku, Fukuoka City, Fukuoka 819-0395, Japan*
[68]*Research Institute for Science and Engineering, Waseda University, Shinjuku, Tokyo 169-8555, Japan*
[69]*Department of Electronic Control Engineering, National Institute of Technology, Nagaoka College, Nagaoka City, Niigata 940-8532, Japan*
[70]*Kavli Institute for the Physics and Mathematics of the Universe (IPMU), Kashiwa City, Chiba 277-8583, Japan*
[71]*Department of Information & Management Systems Engineering, Nagaoka University of Technology, Nagaoka City, Niigata 940-2188, Japan*
[72]*Graduate School of Science and Engineering, Hosei University, Koganei City, Tokyo 184-8584, Japan*
[73]*Faculty of Science, Toho University, Funabashi City, Chiba 274-8510, Japan*
[74]*Max Planck Institute for Gravitational Physics, Potsdam Science Park, Am Mühlenberg 1, D-14476 Potsdam, Germany*
[75]*Faculty of Information Science and Technology, Osaka Institute of Technology, Hirakata City, Osaka 573-0196, Japan*
[76]*National Metrology Institute of Japan, National Institute of Advanced Industrial Science and Technology, Tsukuba City, Ibaraki 305-8568, Japan*
[77]*University of Camerino, via Madonna delle Carderi 9, 62032 Camerino (MC), Italy*
[78]*Istituto Nazionale di Fisica Nucleare, University of Perugia, Perugia 06123, Italy*

[79]*Faculty of Science, Department of Physics, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong, Hong Kong*
[80]*Faculty of Science, Niigata University, Nishi-ku, Niigata City, Niigata 950-2181, Japan*
[81]*Department of Communications, National Defense Academy of Japan, Yokosuka City, Kanagawa 239-8686, Japan*
[82]*Department of Physics, University of Florida, Gainesville, FL 32611, USA*
[83]*Department of Physics and Astronomy, Sejong University, Gwangjin-gu, Seoul 143-747, Korea*
[84]*Department of Physics, Korea University, Seongbuk-gu, Seoul 02841, Korea*
[85]*Department of Astronomy, Beijing Normal University, Beijing 100875, China*
*E-mail: kjun0107@resceu.s.u-tokyo.ac.jp

Received August 8, 2019; Revised March 30, 2020; Accepted March 31, 2020; Published May 28, 2020

We apply independent component analysis (ICA) to real data from a gravitational wave detector for the first time. Specifically, we use the iKAGRA data taken in April 2016, and calculate the correlations between the gravitational wave strain channel and 35 physical environmental channels. Using a couple of seismic channels which are found to be strongly correlated with the strain, we perform ICA. Injecting a sinusoidal continuous signal in the strain channel, we find that ICA recovers correct parameters with enhanced signal-to-noise ratio, which demonstrates the usefulness of this method. Among the two implementations of ICA used here, we find the correlation method yields the optimal results for the case of environmental noise acting on the strain channel linearly.

Subject Index E02, F30

1. Introduction

Though Einstein showed the existence of a gravitational wave solution in his theory of general relativity in 1916, it has taken exactly a century for mankind to succeed in its direct detection. This delay is primarily due to the fact that the gravitational force is exceedingly weak compared with other interactions.

The first detection of a gravitational wave by the advanced Laser Interferometer Gravitational wave Observatory (aLIGO) [1] had a great impact on science and represented the beginning of gravitational wave astronomy. Following aLIGO and advanced Virgo, the large-scale cryogenic gravitational wave telescope (LCGT), now known as KAGRA, has been constructed in Kamioka, Japan [2]. KAGRA will play very important roles in the international network of gravitational wave detection by measuring a number of polarization properties, which is indispensable in proving general relativity [3], and by improving the sky localization of each event significantly [4]. As the first underground and cryogenic detector, it will also provide important information to the third-generation detectors.

Because gravity is the weakest force among the four elementary interactions, gravitational waves have high penetrating power. Therefore, unlike electromagnetic waves, they can propagate without being influenced by the interstellar medium. In the same way, they enable us to see deep inside dense matter, such as the core of neutron stars, and bring information that electromagnetic waves cannot. On the other hand, due to this property, gravitational wave signals tend to be quite small and their detection becomes very difficult. Thus, it is important to develop methods for extraction of these tiny signals. There are a number of methods which extract signal from large amounts of noise, such as matched filtering [5], which yields an optimal result if (and only if) the underlying noise is Gaussian distributed. However, the problem is not so simple, as it is known that non-Gaussian noise exists in real data, which decreases the performance of analysis methods assuming

the Gaussianity of the noise. What is worse, the noise may be mistaken for true signals, increasing the false alarm probability. Thus, it is necessary to deal with non-Gaussianity properly, as stressed in Ref. [6]. Characterization, mitigation, and even subtraction of the noise in gravitational wave detector output have been extensively studied in the literature. The standard methods, including pre-data conditioning (whitening, band-passing), line removal, and χ^2 veto are described well in the overview in Ref. [7]. Many recent works demonstrate the performance of deep neural networks [8–13], but see also Ref. [14].

In this situation, independent component analysis (ICA) [15–17], which separates mixed signal components, occupies a unique position among signal processing methods because it makes use of the non-Gaussianity of signals and noise instead of treating it as an obstacle. ICA has been used in various fields in astronomy, see, e.g., Refs. [18–26]. For example, Ref. [18] demonstrated ICA (EFICA and WASOBI) performance on simulated data mimicking two gravitational wave interferometer outputs. The current paper, on the other hand, demonstrates it using real gravitational wave strain data from the iKAGRA detector and multiple real auxiliary channels that recorded the status of the detector. ICA can separate various components obeying non-Gaussian distributions, so that it can remove (part of) the non-Gaussian noise from strain data that records gravitational wave signals. Then, the strain channel consists of the real signal and (nearly) Gaussian noise. Therefore, ICA can support the conventional matched filter technique as a non-Gaussian noise subtraction scheme. In addition, ICA can be used even in the case when the noise is nonlinearly coupled to the strain channel, as demonstrated in Ref. [27].

In this paper we use the correlation method [27] (or the Gram–Schmidt orthogonalization method in the case of multiple channels in general) and FastICA [31]. The former method, although conceptually different in derivation, has practically the same expression as Wiener filtering [33–38], which has been used to analyze Caltech 40 m and LIGO data, and quite remarkable success was recently reported [38] by using witness sensors, including voltage monitors of the analog electronics for the power main and photodiodes that monitor the beam motion and its size for beam jitter. We report the results of the application of these two different ICA methods to the iKAGRA data and discuss their usefulness in gravitational wave data analysis. The paper is organized as follows. In Sect. 2 we introduce ICA in the simplest case where only one environmental channel is incorporated in the strain channel and review the analytic formulas of the correlation method obtained in our previous paper [27]. Then we extend this method to the case where two different environmental channels are included. We also introduce FastICA, which is formulated in a different way. In Sect. 3 we present our application of ICA to the iKAGRA data with an injected artificial continuous signal. Then, we discuss the results, focusing on the difference between the two methods, in Sect. 4. We argue that for the current setup, where noise measured by the environmental channels affects the strain linearly and additively, what we call the correlation method yields the optimal result. Finally, Sect. 5 is devoted to our conclusions.

2. Independent component analysis

As seen in our previous paper [6], signal detection under non-Gaussian noise is much more involved than the case with Gaussian noise since the optimal statistic has a much more complicated form. ICA is an attractive signal processing method because it makes use of the non-Gaussian nature of the signals [15–17] (see Refs. [28,29] for textbooks). We introduce here two ICA methods for non-Gaussian noise subtraction.

Basically, this method only assumes statistical independence between the signal and the noise, and does not impose any other conditions on their distributions. However, a simpler formulation can be achieved by using physical information from gravitational wave detection, as expressed in Ref. [27], following which, we first formulate the subtraction of non-Gaussian noise in the gravitational wave detection for the case where the noise is linearly coupled to the strain. Then we introduce analytic ICA formulas for this case, which we call the correlation method. Previously, this was two-component analysis in Ref. [27], but we have developed here a multiple-component version for combining different environmental channels.

On the other hand, there is a robust formulation which does not incorporate any information about the system concerned, which is called FastICA [31]. We also introduce this method in this section and apply it in our analysis for comparison.

2.1. Removing non-Gaussian noise

In this paper we consider the following simple problem as a first step to testing the applicability of ICA for gravitational wave detection. Let us consider the case where we have two detector outputs, $x_1(t)$ and $x_2(t)$, where t stands for time. The former is the output from the laser interferometer, namely, the strain channel, and the latter is an environmental channel such as the output of a seismograph. We wish to separate the gravitational wave signal $h(t)$ and non-Gaussian noise $k(t)$ using the data of ${}^t\boldsymbol{x}(t) = (x_1(t), x_2(t))$.

As the simplest case we assume that there is a linear relation between the outputs and the sources:

$$\boldsymbol{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = A\boldsymbol{s}(t), \qquad \boldsymbol{s}(t) = \begin{pmatrix} s_1(t) \\ s_2(t) \end{pmatrix} = \begin{pmatrix} h(t) + n(t) \\ k(t) \end{pmatrix}, \tag{1}$$

where A is assumed to be a time-independent matrix. Since the output of a laser interferometer, of course, suffers from Gaussian noise $n(t)$, we can regard $s_1(t) = h(t) + n(t)$ as the original signal. Note that the non-Gaussian noise $k(t)$ can contain any Gaussian noise as part of it. Thus, we have not added any Gaussian noise to $s_2(t)$ explicitly.

Since the gravitational wave is so weak that it will not affect any environmental meters such as a seismograph, one may set A as

$$A = \begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix}. \tag{2}$$

The aim of ICA is to find a linear transformation

$$\boldsymbol{y} = W\boldsymbol{x} \tag{3}$$

such that two components of the transformed variables $\boldsymbol{y}$ are statistically independent of each other. Here, the distribution of $\boldsymbol{y}$, $p_y(\boldsymbol{y})$, is constructed from the observed distribution function of $\boldsymbol{x}$, $p_x(\boldsymbol{x})$, through the transformation in Eq. (3) as

$$p_y(\boldsymbol{y}) \equiv ||W^{-1}||p_x(\boldsymbol{x}), \tag{4}$$

where $||X||$ denotes the determinant of matrix X. Thanks to the assumption in Eq. (2), the matrix W also takes the form

$$W = \begin{pmatrix} w_{11} & w_{12} \\ 0 & w_{22} \end{pmatrix}. \tag{5}$$

However, since we do not know all the component of A, we attempt to determine W to be A^{-1} in such a way that the components of $\boldsymbol{y}$, $y_1(t)$ and $y_2(t)$, are statistically independent as far as possible. In Ref. [27] this was achieved by using the the Kullback–Leibler divergence [30], which represents distance in the space of statistical distribution functionals. It is defined between two arbitrary probability distribution functions (PDFs), e.g. $p_y(\boldsymbol{y})$ and $q(\boldsymbol{y})$, as

$$D[p_y(\boldsymbol{y}); q(\boldsymbol{y})] = \int p_y(\boldsymbol{y}) \ln \frac{p_y(\boldsymbol{y})}{q(\boldsymbol{y})} dy = E_{p_y}\left[\ln \frac{p_y(\boldsymbol{y})}{q(\boldsymbol{y})}\right]. \tag{6}$$

Here, $E_{p_y}[\cdot]$ denotes an expectation value with respect to a PDF p_y. Then we can obtain mutually independent variables $\boldsymbol{y}$ by minimizing a cost function $L_q(W) \equiv D[p_y(\boldsymbol{y}); q(\boldsymbol{y})]$, where $q(\boldsymbol{y}) = q(y_1)q(y_2)$ is an appropriately chosen distribution function.

The best choice of $q(\boldsymbol{y})$ is obviously the true distribution function of the independent source variables $\boldsymbol{s}$, $r(\boldsymbol{s}) = r_1[s_1(t)]r_2[s_2(t)]$, which is not known a priori. Because $n(t)$ is Gaussian with vanishing mean in this simple setup, its statistical property is entirely characterized by the two-point correlation function $K(t-t') = \langle n(t)n(t')\rangle$. Then, the marginal distribution function of $s_1(t)$ is given by

$$r_1[s_1(t)] = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2\sigma^2}\left(s_1(t) - h(t,\theta)\right)^2\right], \qquad \sigma^2 = K(0), \tag{7}$$

where $h(t,\theta)$ is the actual waveform of the gravitational radiation emitted from some source, where θ collectively denotes the parameters of the source. On the other hand, we do not specify the PDF of $k(t)$, $r_2(s_2)$, except that it is a super-Gaussian distribution such as a Poisson distribution with a larger tail than Gaussian. We show, however, that we can obtain the matrix W easily for our particular problem with $a_{21} = w_{21} = 0$, as we see below.

2.2. *Correlation method*

From now on we replace the ensemble average $E[\cdot]$ by the temporal average of observed values of $\boldsymbol{x}$, which we denote by brackets. For the true distribution $r(\boldsymbol{y})$, minimization of the cost function $L_r(W)$ results in decorrelating y_1 and y_2 [27], i.e. $\langle y_1(t)y_2(t)\rangle = 0$.

From

$$\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = \begin{pmatrix} w_{11} & w_{12} \\ 0 & w_{22} \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} w_{11}x_1(t) + w_{12}x_2(t) \\ w_{22}x_2(t) \end{pmatrix}, \tag{8}$$

it is equivalent to requiring

$$\langle y_1(t)x_2(t)\rangle = w_{11}\langle x_1(t)x_2(t)\rangle + w_{12}\langle x_2^2(t)\rangle = 0. \tag{9}$$

We therefore obtain

$$w_{12} = -\frac{\langle x_1 x_2 \rangle}{\langle x_2^2 \rangle} w_{11}. \tag{10}$$

Since ICA does not uniquely determine the overall factor of $\boldsymbol{y}$, this relation suffices for our purpose to determine y_1. These are what we calculated in our previous paper [27] using the Kullback–Leibler divergence.

Here we develop a multiple-component method for further analysis; we apply it in Sect. 3.2.2. For three components, $\boldsymbol{y}(t)$ and $\boldsymbol{x}(t)$ become

$$\begin{pmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{pmatrix} = \begin{pmatrix} w_{11} & w_{12} & w_{13} \\ 0 & w_{22} & w_{23} \\ 0 & w_{32} & w_{33} \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} = \begin{pmatrix} w_{11}x_1(t) + w_{12}x_2(t) + w_{13}x_3(t) \\ w_{22}x_2(t) + w_{23}x_3(t) \\ w_{32}x_2(t) + w_{33}x_3(t) \end{pmatrix} \tag{11}$$

and

$$\begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} s_1(t) \\ s_2(t) \\ s_3(t) \end{pmatrix} = \begin{pmatrix} a_{11}s_1(t) + a_{12}s_2(t) + a_{13}s_3(t) \\ a_{22}s_2(t) + a_{23}s_3(t) \\ a_{32}s_2(t) + a_{33}s_3(t) \end{pmatrix}. \tag{12}$$

In this case also, the minimization of the cost function results in decorrelating $\boldsymbol{y}$, $\langle y_1 y_2 \rangle = \langle y_2 y_3 \rangle = \langle y_3 y_1 \rangle = 0$. This is achieved by an analog of the Gram-âŁ“Schmidt process, which is a method for orthonormalizing a set of vectors, and it can be extended to the case where there are more than three components.

Because of the gauge degree of freedom, we can take $w_{32} = 0$ without loss of generality and choose

$$y_3(t) = \tilde{x}_3(t) \equiv \frac{x_3(t)}{\sqrt{\langle x_3^2 \rangle}}. \tag{13}$$

We first require $\langle y_2(t) y_3(t) \rangle = \langle y_2(t) x_3(t) \rangle = 0$. This gives following relation:

$$w_{23} = -\frac{\langle x_2 x_3 \rangle}{\langle x_3^2 \rangle} w_{22}. \tag{14}$$

Based on this, we can choose

$$y_2(t) = \tilde{x}_2(t) \equiv \frac{x_2'(t)}{\sqrt{\langle x_2'^2 \rangle}}, \qquad x_2'(t) \equiv x_2(t) - \frac{\langle x_2 x_3 \rangle}{\langle x_3^2 \rangle} x_3(t). \tag{15}$$

If we take

$$y_1(t) = x_1(t) - \langle x_1 \tilde{x}_2 \rangle \tilde{x}_2(t) - \langle x_1 \tilde{x}_3 \rangle \tilde{x}_3(t), \tag{16}$$

$\langle y_1(t) y_2(t) \rangle = \langle y_2(t) y_3(t) \rangle = \langle y_3(t) y_1(t) \rangle = 0$ is satisfied. Note that Eq. (16) is symmetrical with respect to permutation of $x_2(t)$ and $x_3(t)$.

Thus we can observe that the ICA correlation method shown here is equivalent to instantaneous Wiener filtering,[1] and this is due to the particular character of our problem that only the strain channel is sensitive to the gravitational wave signal, with $a_{i1} = 0$ $(i \neq 1)$ in our linear model.

2.3. *FastICA method*

Next, we introduce another method to obtain the matrix W, called FastICA [31], which can be easily implemented even when $\boldsymbol{x}(t) = A\boldsymbol{s}(t)$ has more than two components. Note that this method can be applied to various cases of signal separation other than the case formulated in Sect. 2.1.

In this method, assuming that each component, $s_i(t)$, of the source vector $\boldsymbol{s}(t)$ is properly normalized with vanishing mean, we first apply whitening to the detector outputs $\boldsymbol{x}(t)$ and take the dispersion of each source $s_i(t)$ to be unity. This is achieved in the following way. First, let the normalized eigenvector and corresponding eigenvalue of a matrix $\langle \boldsymbol{x}\, {}^t\boldsymbol{x}\rangle$ be $\boldsymbol{c}_i$ and λ_i, respectively $(i = 1, 2, \ldots)$, and define a matrix Γ by $\Gamma = (\boldsymbol{c}_1, \boldsymbol{c}_2, \boldsymbol{c}_3, \ldots)$, and $\Lambda^{-1/2}$ by $\Lambda^{-1/2} = \mathrm{diag}(\lambda_1^{-1/2}, \lambda_2^{-1/2}, \ldots)$. Then the whitened variable $\tilde{\boldsymbol{x}}(t)$ is defined by

$$\tilde{\boldsymbol{x}}(t) = \Lambda^{-1/2}\, {}^t\Gamma \boldsymbol{x} = \Lambda^{-1/2}\, {}^t\Gamma A\boldsymbol{s} \equiv \tilde{A}\boldsymbol{s}, \tag{17}$$

which satisfies

$$\langle \tilde{\boldsymbol{x}}(t)\, {}^t\tilde{\boldsymbol{x}}(t)\rangle = \langle \tilde{A}\boldsymbol{s}\, {}^t(\tilde{A}\boldsymbol{s})\rangle = \tilde{A}\langle \boldsymbol{s}\, {}^t\boldsymbol{s}\rangle {}^t\tilde{A} = \tilde{A}\, {}^t\tilde{A} = E. \tag{18}$$

Here we have used the statistical independence of each component of the normalized source term s_i. This means that the matrix $\tilde{A}$ is an orthogonal matrix and that W may be identified with ${}^t\tilde{A}$ for whitened output data $\tilde{\boldsymbol{x}}$. Thus, we may restrict W to be an orthogonal matrix, too, after appropriate whitening.[2]

We choose $q(\boldsymbol{y})$ as a product of marginal distributions here,

$$q(\boldsymbol{y}) = \tilde{p}_y(\boldsymbol{y}) \equiv \prod_i \tilde{p}_i(y_i), \qquad \tilde{p}_i(y_i) = \int p_y(\boldsymbol{y}) dy_1 \cdots dy_{i-1} dy_{i+1} \cdots, \tag{19}$$

since $p_y(\boldsymbol{y}) = \tilde{p}_y(\boldsymbol{y})$ is the condition for statistical independence of the variables $\boldsymbol{y}$. Then, the cost function defined in terms of the Kullback–Leibler divergence reads

$$L_{\tilde{p}}(W) = D[p_y(\boldsymbol{y}); \tilde{p}_y(\boldsymbol{y})] = -H[\boldsymbol{x}] - \ln ||W|| + \sum_i H_i[y_i], \tag{20}$$

where $H[\boldsymbol{x}] \equiv -\int d\boldsymbol{x} p_x(\boldsymbol{x}) \ln p_x(\boldsymbol{x})$ is the entropy of the distribution of $\boldsymbol{x}$, and $H_i[y_i] \equiv -\int dy_i \tilde{p}_i(y_i) \ln \tilde{p}_i(y_i)$ is the entropy of the marginal distribution of y_i. When W is an orthonormal matrix, only the last term matters in determining it. Hence, minimization of the cost function for W is achieved by minimizing the entropy of the marginal distribution of each variable. This is the spirit of the FastICA method. It has been proposed to maximize the negentropy defined by

$$J[y_i] \equiv H[v] - H[y_i], \tag{21}$$

[1] The Wiener filtering adopted in Refs. [33–37] takes into account the time delay in transfer functions. We can easily incorporate this in our analysis, too, as already demonstrated in Ref. [27].

[2] Note that this procedure is also called sphering, and has nothing to do with the whitening of the strain data in the frequency domain.

which is a positive semi-definite quantity, instead of the entropy itself. Here, ν is a random Gaussian variable with vanishing mean and unit variance.

In order to achieve easier implementation of the method, however, we minimize a simpler cost function $L(\boldsymbol{w}_i)$ for each row vector $\boldsymbol{w}_i$ constituting the matrix W as $W \equiv (\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots)$. Since W is an orthogonal matrix now, we find $|\boldsymbol{w}_i|^2 = 1$, so the cost function may be defined as

$$L(\boldsymbol{w}_i) = \{E[G(y_i)] - E[G(\nu)]\}^2 - \beta \left[|\boldsymbol{w}_i|^2 - 1\right], \tag{22}$$

where G is an appropriate nonquadratic function and β is a Lagrange multiplier. Minimization of Eq. (22) corresponds to solving the following equation:

$$E[\tilde{\boldsymbol{x}} g({}^t\boldsymbol{w}_i \tilde{\boldsymbol{x}})] - \beta \boldsymbol{w}_i = 0, \tag{23}$$

where $g(y) = G'(y)$. FastICA solves for this equation, starting from an arbitrary initial choice of $\boldsymbol{w}_i$, in terms of the Newton method.

3. Analysis of iKAGRA data

The initial engineering run of KAGRA without the cryogenic system was performed in March and April 2016 [32]. From the results of all the time series data that we analyzed, we report those of two particular datasets each of length 224 s. One starts from 20:15:11 UTC on April 14, 2016, and the other starts from 01:01:35 UTC on April 17, 2016. For each dataset we calculated Pearson's correlation between the strain channel and each of 35 physical environmental monitor (PEM) channels. We found that almost all these channels in the latter (former) dataset strongly (weakly) correlated with the strain channel. We call the latter (former) the strongly (weakly) correlated data. The amplitude spectrum density (ASD) of the strain channel for each dataset is depicted in Fig. 1.

We chose two channels which showed large correlation with the strain channel for each dataset; these are listed in Table 1. For both datasets, channel 4724ch had the largest correlation with the strain. This channel is the output of the seismograph that observes vertical vibration installed at the end of the X arm, 4774ch and 4823ch are the outputs of the seismographs installed at the end of the Y arm, and they observe horizontal vibration orthogonal to each other.

We produced mock strain data by injecting sinusoidal continuous waves,

$$s(t) = A \sin(2\pi f t), \tag{24}$$

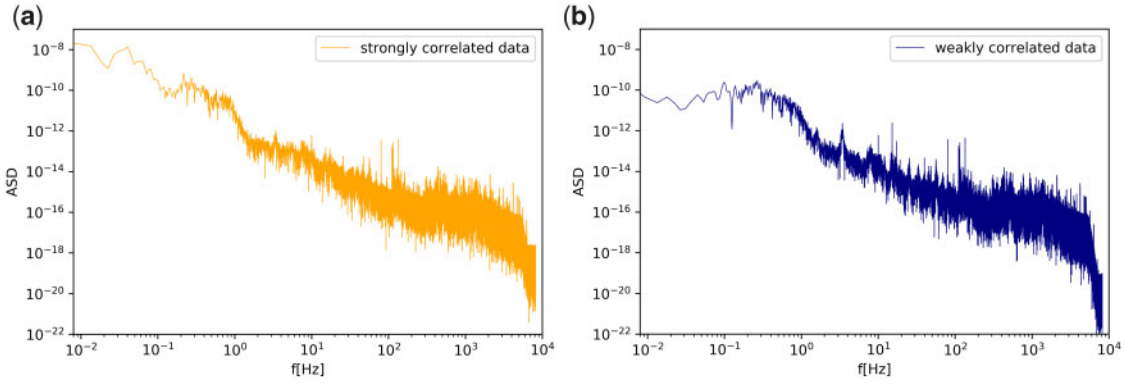

Fig. 1. Amplitude spectrum densities of the strain channels for two datasets. The ASD below 0.1 Hz for the strongly correlated data becomes much larger than that of the weakly correlated data. This means that the strongly correlated data is contaminated by seismic noise at lower frequencies.

Table 1. Correlation between PEM channels and strain.

Dataset	Channel	Correlation coefficient
Strongly correlated	`PEM-EX_SEIS_Z_SENSINF_OUT16` (4724ch)	−0.6409
	`PEM-EY_SEIS_WE_SENSINF_OUT16` (4823ch)	0.5892
Weakly correlated	`PEM-EX_SEIS_Z_SENSINF_OUT16` (4724ch)	0.3078
	`PEM-EY_SEIS_NS_SENSINF_OUT16` (4774ch)	−0.2312

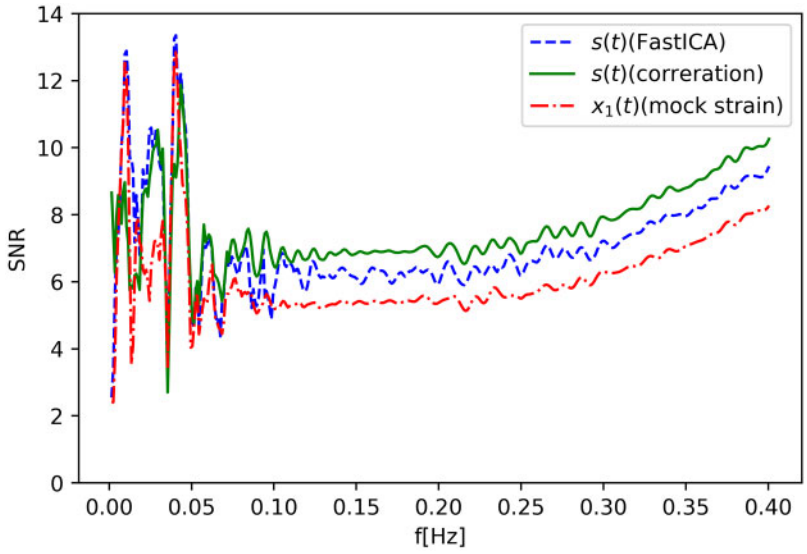

Fig. 2. Signal-to-noise ratio for varying f with and without ICA using 4724ch for the strongly correlated dataset. The red line corresponds to the raw mock strain, while the green and blue lines are noise-removed data using the correlation method and FastICA, respectively.

into the strain channel, and applied the two ICA methods introduced in the previous section to this mock data and those environmental channels.

We utilized the Python implementation of FastICA from `scikit-learn`.[3] We found that the results often depend on the initial conditions where the Newton method is started. To mitigate this, we generated at most 30 realizations in parallel and chose the one giving the highest signal-to-noise ratio (SNR).

3.1. Global performance

First, we analyze how the SNR changes before and after noise separation by ICA for mock data with varying frequencies f. We performed matched filter (MF) analysis on both the raw mock strain data and the noise-removed data in terms of the two ICA methods using 4724ch as an environmental channel. For various f of the injected signal in Eq. (24), we calculated the SNR by applying an MF with the same frequency as the injected signal. We simultaneously plotted the results against the data before and after ICA to assess the global performance of ICA. For the strongly correlated data, the results are shown in Fig. 2.

[3] https://scikit-learn.org/stable/.

In this dataset, the strain had a larger amplitude than in the other dataset, and we set $A = 9 \times 10^{-10}$. As one can see from Fig. 2, the SNRs are homogeneously enhanced by ICA for $f \gtrsim 0.1$ Hz. The correlation method enhances the SNR more than FastICA. However, there are anomalous peaks at frequencies of 0.01 Hz and 0.04 Hz. As shown in Fig. 1(a), even in the absence of injection the strain channel has a large amplitude at these frequencies, which is predominantly contributed by seismic noise. We also found that their oscillation phases are more or less stable during the time period we analyzed. Such noise is difficult to distinguish from our sinusoidal signal waveform and hence yields a large SNR for the mock strain, as shown in Fig. 2. This, however, indicates that by removing the contribution of the noise, the SNR can possibly be reduced rather than enhanced, provided the injected signal is moderate. This is actually realized in the analysis based on the correlation method, as seen in Fig. 2.

On the other hand, in the case of FastICA, the reduction of SNR is not seen. This is solely due to our implementation, which tries to increase the SNR as much as possible, as mentioned before. In that sense, around the 0.01 Hz and 0.04 Hz peaks, the blue line in Fig. 2 corresponds to the SNR of the separated noise.

Apart from these low frequencies contaminated by seismic noise, we find that ICA improves SNR significantly throughout the entire frequency range with $f \gtrsim 0.1$ Hz. However, based on these considerations, it is deduced that ICA works even near the peak due to seismic noise.

For the weakly correlated data, the results are shown in Fig. 3. The amplitude of the strain in this time period is moderate, and we set $A = 3 \times 10^{-11}$. As is seen in Fig. 3, the SNR of the data with ICA is higher than the mock data in several frequency ranges. Comparing FastICA with the correlation method, the correlation method has fewer frequencies where the SNR falls below that of the mock data.

As for the weakly correlated data, 4774ch had the second highest correlation with the strain. If we use 4774ch instead of 4724ch as the environmental data, the result changes as shown in Fig. 4. Compared with 4724ch (Fig. 3), the frequency region where the SNR rises is different. As a whole, the improvement of SNR is less significant, which is a natural result considering that the correlation coefficient of 4774ch is smaller than that of 4724ch.

3.2. Parameter estimation for strongly correlated data

3.2.1. Two-channel ICA

Next, we perform parameter estimation using the strongly correlated data to examine whether ICA can recover the correct parameters of injected signals. We injected the sinusoidal waveform in Eq. (24) with $f = 0.125$ Hz and $A = 1.3 \times 10^{-9}$. We applied MF analysis to search for the frequency with the highest SNR, which corresponds to the maximum likelihood estimation of the parameter. We compare how the result of parameter estimation changes before and after ICA and by how much the SNR changes.

Figure 5 depicts the SNR before and after applying ICA. In this case, we can see the effect of seismic noise directly. By ICA with 4724ch, the SNR at $f \sim 0.01$ Hz is reduced and that at the injected frequency $f = 0.125$ Hz is successfully enhanced. From this result, we deduce that 4724ch is highly correlated to the 0.01 Hz peak. On the other hand, the peak of 0.04 Hz is still higher, which turned out to be correlated to 4823ch, which had the second largest correlation with the strain, as we will see below.

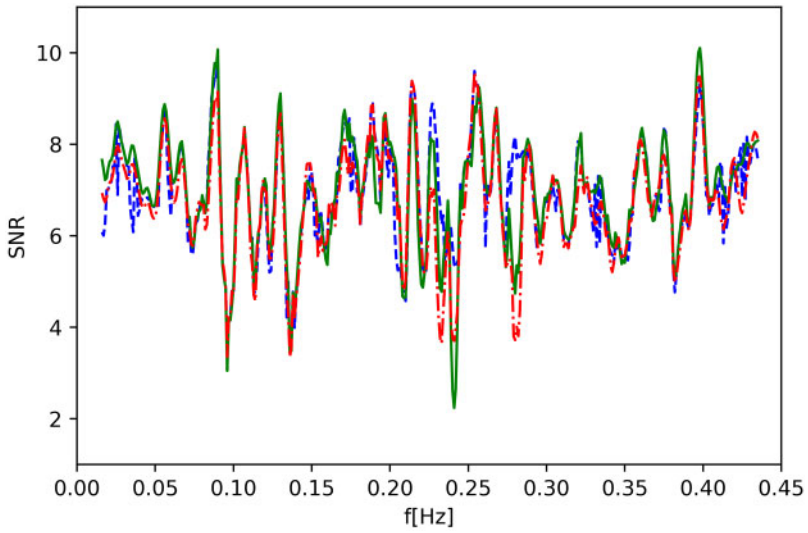

Fig. 3. As Fig. 2 but for the weakly correlated dataset.

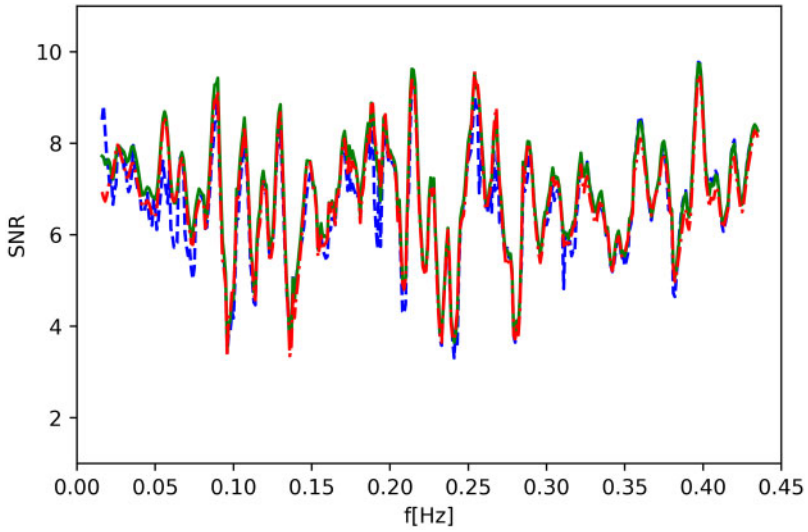

Fig. 4. As Fig. 3 but using 4774ch.

3.2.2. Multiple-channel ICA

As seen in Sect. 3.1, the correlation method shows more stable performance than FastICA, although it is much simpler. This method can be generalized to multi-channel analysis. As a first step to this, we investigate here the effectiveness of three-component analysis, which we developed in Sect. 2.2, including two PEM channels which strongly correlated with the strain. For this purpose we have used mock data including the same signal waveform as in the previous subsection, and applied the three-component correlation method to this mock data on 4724ch and 4823ch.

The result is shown in Fig. 6, which also shows the results of two-component analysis in which we used 4724ch and 4823ch respectively. The green and black lines correspond to the cases where noises are removed using one PEM channel. While the 0.01Hz peak was reduced by using 4724ch, the 0.04Hz peak was reduced by using 4823ch. However, both peaks cannot be reduced when we use only one PEM channel. The data with ICA using two PEM channels (cyan line) has much higher

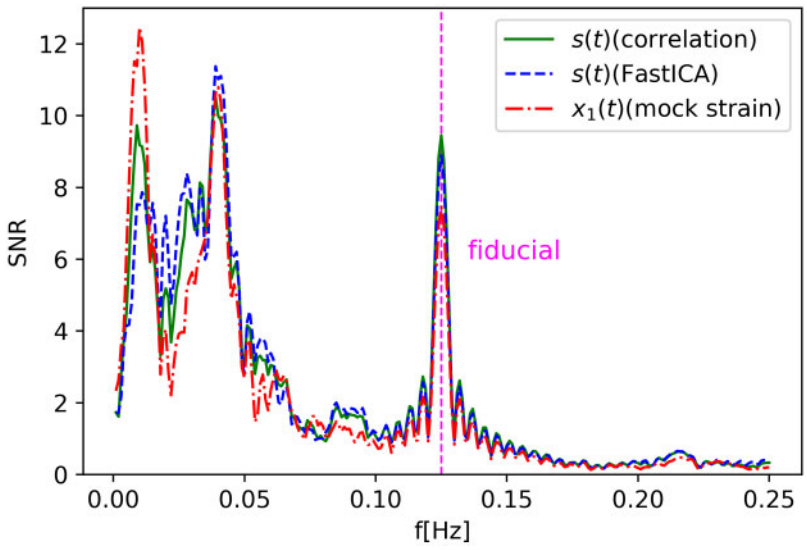

Fig. 5. Parameter estimation with fiducial frequency $f = 0.125$ Hz. The lines are the same as in Fig. 2.

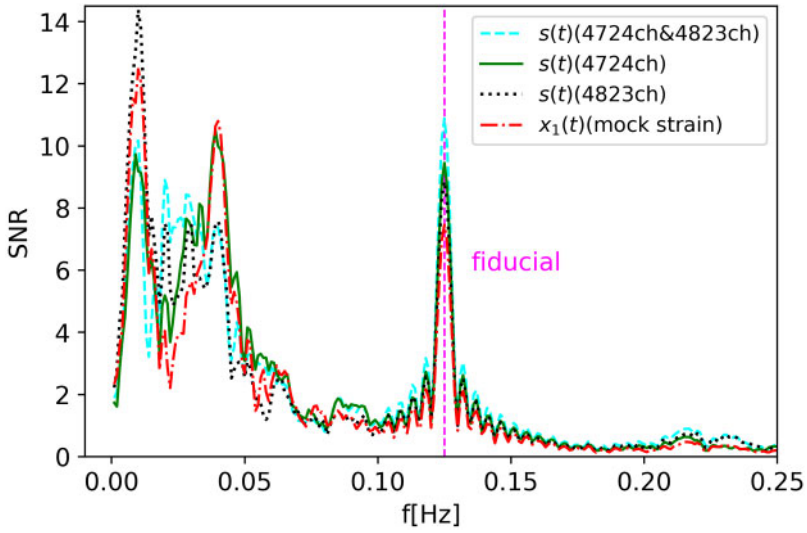

Fig. 6. Parameter estimation with multiple-channel ICA (correlation method).

SNR than the data with ICA using only one PEM channel. In addition, we successfully reduced both 0.01Hz peak and 0.04Hz peak. This result suggests that by combining many environmental channels we can effectively remove noises with various characteristic frequencies.

As explained in Sect. 2.3, FastICA can be easily implemented even when there are more than two components. We applied FastICA to the mock data on 4724ch and 4823ch simultaneously. Here, the mock data included the same sinusoidal signal as in the previous section. The result is shown in Fig. 7.

As compared to Fig. 5, the SNR at the fiducial frequency is much higher than the case where only 4724ch was used. In addition, its value is close to that for the three-component correlation method (10.60 for FastICA, 10.89 for the correlation method). This result suggests that the use of multiple environmental channels can also enhance the effect of FastICA noise separation. However,

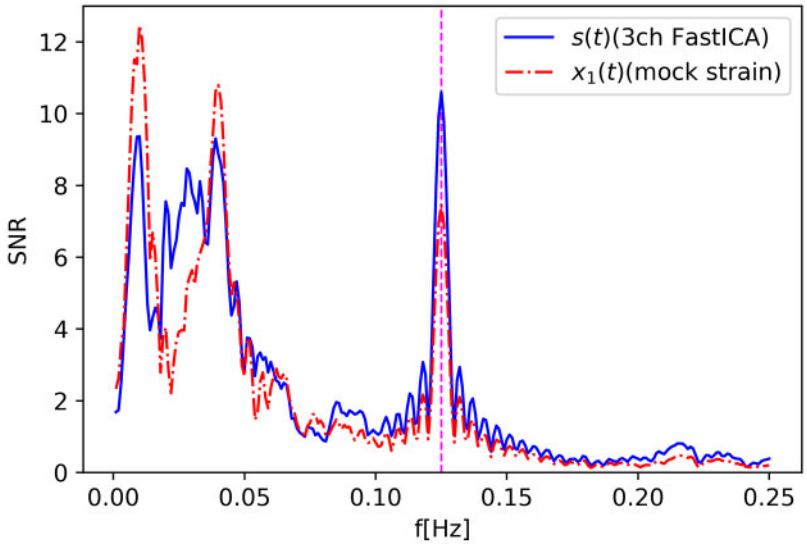

Fig. 7. Parameter estimation with multiple-channel ICA (FastICA).

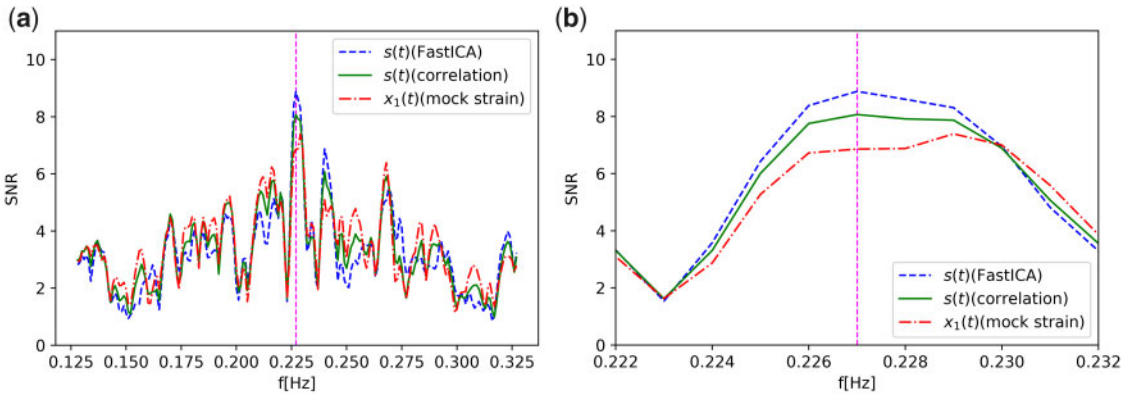

Fig. 8. Parameter estimation for the weakly correlated data with the fiducial frequency $f = 0.227\,\mathrm{Hz}$.

compared to the three-component correlation method, we may have to make several trials of three-channel FastICA to obtain the best result. This indicates that the correlation method is more effective than FastICA for this dataset.

3.3. Parameter estimation for weakly correlated data

We also perform parameter estimation for weakly correlated data. Here, we used 4724ch as an environmental channel. From Fig. 3, ICA using 4724ch is most effective for $f = 0.227\,\mathrm{Hz}$ with this dataset. We injected a sinusoidal wave signal with $f = 0.227\,\mathrm{Hz}$ and $A = 3 \times 10^{-11}$. Again, we applied MF to search for the frequency with the highest SNR. The result is depicted in Fig. 8, where the red line represents the SNR calculated with the raw mock strain, and the green and blue lines correspond to the strain with noise removed by the correlation method and FastICA, respectively. An enlarged figure of the fiducial ($f = 0.227\,\mathrm{Hz}$) area is shown in Fig. 8(b). As one can see, in the case of the raw mock strain, the position of the SNR peak deviates from the fiducial one. On the other hand, after applying ICA, the SNR is increased and the peak is found at the correct frequency.

Next, we applied multiple-channel ICA to this data using 4724ch and 4774ch. Here, we used the correlation method. Figure 9 depicts the results of the analysis. An enlarged figure of the fiducial

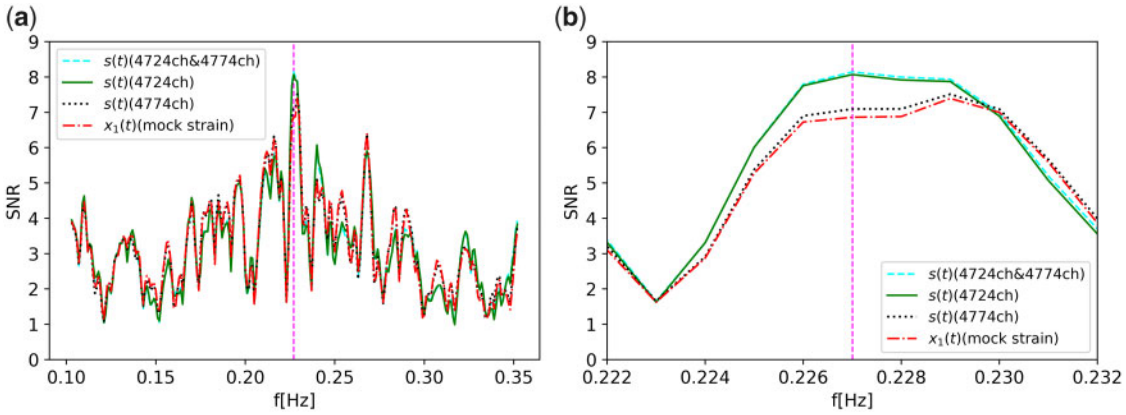

Fig. 9. Parameter estimation with multiple-channel ICA (correlation method).

area is shown in Fig. 9(b). The green and black lines correspond to the data with ICA using one PEM channel. When using only 4774ch, the enhancement of the SNR is small and the SNR peak still deviates from the fiducial frequency. However, the data with ICA using two channels has a slightly higher SNR at the correct frequency than the data with ICA using only 4724ch. This result for the weakly correlated data also supports our expectation that the effect of ICA can be enhanced by combining many environmental channels.

4. Discussion

In Sect. 3 we showed the performance of ICA as a method of non-Gaussian noise subtraction for gravitational wave data. Both ICA methods, namely the correlation method and FastICA, subtracted a portion of the seismic noise. However, the correlation method shows better performance than FastICA in most cases. In this section we consider the reason why this difference appears.

We use the same notation as in Sect. 2, $x_1(t)$ being the strain channel and $x_i(t)$ $(i = 2, \ldots, n)$ the other environmental channels. As discussed in Sect. 2.1, the data from these channels can be written in the following form:

$$x_1(t) = h(t) + n(t) + \sum_{j=2}^{n} a_{1j} s_j(t),$$
$$x_i(t) = \sum_{j=2}^{n} a_{ij} s_j(t). \tag{25}$$

Here, $s_i(t)$ $(i = 2, \ldots, n)$ are environmental noise that can be measured by the PEM channels $x_i(t)$, and $n(t)$ collectively represents the noise of the strain channel to which these PEM channels are insensitive. Let us transform x_1 as

$$\tilde{x}_1 = x_1 + \sum_{j=2}^{n} b_{1j} x_j \tag{26}$$

in order to satisfy $\langle \tilde{x}_1 x_i \rangle = 0$ $(i = 2, \ldots, n)$. This condition can be expanded as

$$\langle \tilde{x}_1(t) x_i(t) \rangle = \left\langle \left(h(t) + \sum_{j=2}^{n} a_{1j} s_j(t) + \sum_{j=2}^{n} b_{1j} \sum_{l=2}^{n} a_{jl} s_l(t) \right) \sum_{k=2}^{n} a_{ik} s_k(t) \right\rangle$$

$$= \sum_{j=2}^{n} a_{1j}a_{ij} + \sum_{j=2}^{n} b_{1j} \sum_{k=2}^{n} a_{jk}a_{ik} = 0, \tag{27}$$

where we have used $\langle s_i s_j \rangle = \delta_{ij}$ as in Sect. 2.3. From this equation we obtain

$$b_{1j} = \sum_{i=2}^{n} a_{1i}a_{ij}^{-1}, \tag{28}$$

with $j = 2, \ldots, n$. Note that a_{ij}^{-1} is the inverse matrix of a_{ij} $(i, j = 2, \ldots, n)$, which is an $(n-1)\times(n-1)$ partial matrix of the mixing matrix $A = (a_{ij})_{1\le i,j\le n}$. By substituting this into Eq. (26), we obtain

$$\begin{aligned}
\tilde{x}_1 &= x_1 + \sum_{i=2}^{n} b_{1i}x_i \\
&= h(t) + n(t) + \sum_{i=2}^{n} a_{1i}s_i(t) + \sum_{i=2}^{n}\sum_{j=2}^{n}\sum_{l=2}^{n} a_{1j}a_{ji}^{-1}a_{il}s_l(t) \\
&= h(t) + n(t) + \sum_{i=2}^{n} a_{1i}s_i(t) - \sum_{j=2}^{n}\sum_{l=2}^{n} a_{1j}\delta_{jl}s_l(t) \\
&= h(t) + n(t) + \sum_{i=2}^{n} a_{1i}s_i(t) - \sum_{j=2}^{n} a_{1j}s_j(t) = h(t) + n(t),
\end{aligned} \tag{29}$$

which shows that all environmental noise s_i, measurable by the PEM channels, are removed from $\tilde{x}_1$ just by imposing $\langle \tilde{x}_1 x_i \rangle = 0$ $(i = 2, \ldots, n)$. In other words, when we consider auxiliary channels which are not sensitive to gravitational waves, and their target noise affects the strain linearly and additively, we can obtain the independent component $h(t)$ by the transformation in Eq. (26) which eliminates the two-point correlation between the strain and those channels. Although we have not given a concrete expression for the transformation in Eq. (26), $\langle \tilde{x}_1 x_i \rangle = 0$ is naturally achieved by the correlation method, which is analogous to Gram–Schmidt orthogonalization. Thus, we find that the correlation method is the optimal filter for linearly coupled noise with $a_{i1} = 0$.

On the other hand, FastICA maximizes negentropy after the whitening, which makes $\langle \tilde{x}_1 x_i \rangle = 0$ without using the condition $a_{i1} = 0$ $(i \neq 1)$. Since we do not recover this property in general even after maximizing the negentropy, FastICA tends to show less enhancement of SNR than the optimal correlation method for most cases in our analysis. This illustrates the importance of incorporating the characteristic features of the system as much as possible before applying ICA.

However, the above discussion is only the case where linearly coupled noise with $a_{i1} = 0$ is concerned. For real observational data there will be much more complicated mixing such as nonlinear coupling of the noise, and we might need a formulation of ICA which treats general mixing of signals. In that sense, it is noteworthy that FastICA, which is formulated without any assumption like $a_{i1} = 0$, also shows enhancement of the SNR and improvement of the performance with multi-environmental channels to some extent.

5. Conclusions

We have demonstrated the usefulness of ICA in gravitational wave data analysis in application to the iKAGRA strain and environmental channels. Assuming continuous waves as input signals, we

have shown that ICA can enhance SNR, in particular when the strain channel has a large correlation with the environmental ones. Moreover, we have shown that ICA can correctly recover the input frequencies in parameter estimation. We have also found that combining multiple environmental channels can enhance the effect of ICA to improve SNR.

There are, however, a number of limitations in the analysis presented here because the iKAGRA data contains more low-frequency modes than wanted due to the simplified vibration isolation system compared with the full design specification which will be realized with bKAGRA [39], and the iKAGRA configuration was not equipped with environmental monitors that measure hectohertz frequencies. Hence, we had to concentrate on relatively lower frequency components as the first step of application of ICA to real data analysis of laser interferometers.

Another limitation is that we have restricted to the case where all the environmental noise that can be measured by the PEM channels under consideration act on the strain channel linearly and additively, without incorporating nonlinear couplings. In this particular situation, we have shown that the Gram–Schmidt decorrelation approach, or the instantaneous Wiener filtering which we dubbed the correlation method, gives the optimal result of environmental noise removal as an implementation of ICA. However, ICA can be used even in the case where noise of different origin is nonlinearly coupled to affect the strain channel as demonstrated in Ref. [27]. This is one of the merits of ICA that is absent from other methods. We could not perform such an analysis here due to the limitation of available PEM channels. We plan to return to this issue when the full cryogenic configuration of bKAGRA starts operation with more PEM channels.

Acknowledgements

This work was supported by MEXT, thre Japan Society for the Promotion of Science (JSPS) Leading-edge Research Infrastructure Program, JSPS Grant-in-Aid for Specially Promoted Research 26000005, JSPS Grant-in-Aid for Scientific Research on Innovative Areas 2905 nos. JP17H06358, JP17H06361, and JP17H06364, JSPS Core-to-Core Program A, Advanced Research Networks, JSPS Grant-in-Aid for Scientific Research (S) 17H06133, the joint research program of the Institute for Cosmic Ray Research, University of Tokyo in Japan, National Research Foundation (NRF) and Computing Infrastructure Project of KISTI-GSDC in Korea, Academia Sinica (AS), AS Grid Center (ASGC), and the Ministry of Science and Technology (MoST) in Taiwan under grants including AS-CDA-105-M06, the LIGO project, and the Virgo project. This paper carries JGW Document Number JGW-P1910218. Jun'ya Kume is supported by a research program of the Leading Graduate Course for Frontiers of Mathematical Sciences and Physics (FMSP). This work was partially supported by JSPS KAKENHI Grant-in-Aid for Scientific Research No. 15H02082 (Jun'ichi Yokoyama, Yosuke Itoh, and Toyokazu Sekiguchi).

References

[1] B.P. Abbott et al. [LIGO Scientific Collaboration and Virgo Collaboration], Phys. Rev. Lett. **116**, 061102 (2016).
[2] Y. Aso, Y. Michimura, K. Somiya, M. Ando, O. Miyakawa, T. Sekiguchi, D. Tatsumi, and H. Yamamoto [KAGRA Collaboration], Phys. Rev. D **88**, 043007 (2013).
[3] Y. Hagihara, N. Era, D. Iikawa, and H. Asada, Phys. Rev. D **98**, 064035 (2018) [arXiv:1807.07234 [gr-qc]] [Search INSPIRE].
[4] B. P. Abbott et al. [KAGRA and LIGO Scientific and VIRGO Collaborations], Living Rev. Rel. **21**, 3 (2018) [arXiv:1304.0670 [gr-qc]] [Search INSPIRE].
[5] L. S. Finn, Phys. Rev. D **46**, 5236 (1992).
[6] J. Yokoyama, Proc. Japan Acad. B **90**, 422 (2014).
[7] B. P. Abbott et al. [LIGO Scientific Collaboration and Virgo Collaboration] Class. Quantum Grav. **37**, 055002 (2020) [arXiv:1908.11170 [gr-qc]] [Search INSPIRE].
[8] M. Zevin et al., Class. Quantum Grav. **34**, 064003 (2017) [arXiv:1611.04596 [gr-qc]] [Search INSPIRE].

[9] M. Razzano and E. Cuoco, Class. Quantum Grav. **35**, 095016 (2018).
[10] D. George, H. Shen, and E. A. Huerta, Phys. Rev. D **97**, 101501(R) (2018).
[11] J. Powell, Class. Quantum Grav. **35**, 155017 (2018).
[12] C. Pankow et al., Phys. Rev. D **98**, 084016 (2018) [arXiv:1808.03619 [gr-qc]] [Search INSPIRE].
[13] S. Coughlin et al., Phys. Rev. D **99**, 082002 (2019) [arXiv:1903.04058 [astro-ph.IM]] [Search INSPIRE].
[14] B. Zackay, T. Venumadhav, J. Roulet, L. Dai, and M. Zaldarriaga, arXiv:1908.05644 [astro-ph.IM] [Search INSPIRE].
[15] C. Jutten and J. Herault, Sig. Proc. **24**, 1 (1991).
[16] P. Common, Sig. Proc. **36**, 287 (1994).
[17] S.-I. Amari and J.-F. Cardoso, IEEE Trans. Sig. Proc. **45**, 2692 (1997).
[18] R. De Rosa, L. A. Forte, F. Garufi, and L. Milano, Phys. Rev. D **85**, 042001 (2012).
[19] S. M. Leach et al., Astron. Astrophys. **491**, 597 (2008) [arXiv:0805.0269 [astro-ph]] [Search INSPIRE].
[20] K. Ichiki, R. Kaji, H. Yamamoto, T. T. Takeuchi, and Y. Fukui, Astrophys. J. **780**, 13 (2014) [arXiv:1303.0075 [astro-ph.CO]] [Search INSPIRE].
[21] M. Querejeta et al., Astrophys. J. Supp. **219**, 5 (2015).
[22] D. Alonso, P. Bull, P. G. Ferreira, and M. G. Santos, Mon. Not. Roy. Astron. Soc. **447**, 400 (2015) [arXiv:1409.8667 [astro-ph.CO]] [Search INSPIRE].
[23] G. Morello, I. P. Waldmann, G. Tinetti, I. D. Howarth, G. Micela, and F. Allard. Astrophys. J. **802**, 117 (2015).
[24] J. G. Ingalls et al., Astron. J. **152**, 44 (2016).
[25] G. Morello, I. P. Waldmann, and G. Tinetti. Astrophys. J. **820**, 86 (2016).
[26] Y. D. Hezaveh, L. Perreault Levasseur, and P. J. Marshall, Nature **548**, 555 (2017).
[27] S. Morisaki, J. Yokoyama, K. Eda, and Y. Itoh, Proc. Japan Acad. B **92**, 336 (2016) [arXiv:1605.01983 [gr-qc]] [Search INSPIRE].
[28] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis* (John Wiley, New York, 2001).
[29] A. Cichocki and S. Amari, *Adaptive Blind Signal and Image Processing* (John Wiley, New York, 2002).
[30] S. Kullback and R. A. Leibler, Ann. Math. Stat. **22**, 79 (1951).
[31] A. Hyvarinen, IEEE Trans. Neural Net. **10**, 626 (1999).
[32] T. Akutsu et al. [KAGRA Collaboration], Prog. Theor. Exp. Phys. **2018**, 013F01 (2018) [arXiv:1712.00148 [gr-qc]] [Search INSPIRE].
[33] B. Allen, W. S. Hua, A. C. Ottewill, arXiv:gr-qc/9909083 [Search INSPIRE].
[34] J. C. Driggers, M. Evans, K. Pepper, and R. Adhikari, Rev. Sci. Instrum. **83**, 024501 (2012) [arXiv:1112.2224 [gr-qc]] [Search INSPIRE].
[35] R. DeRosa, J. C. Driggers, D. Atkinson, H. Miao, V. Frolov, M. Landry, J. A. Giaime, and R. X. Adhikari, Class. Quantum Grav. **29**, 215008 (2012).
[36] G. D. Meadors, K. Kawabe, and K. Riles, Class. Quantum Grav. **31**, 105014 (2014) [arXiv:1311.6835 [astro-ph.IM]] [Search INSPIRE].
[37] V. Tiwari et al., Class. Quantum Grav. **32**, 165014 (2015) [arXiv:1503.07476 [gr-qc]] [Search INSPIRE].
[38] J. C. Driggers et al. [LIGO Scientific Collaboration], Phys. Rev. D **99**, 042001 (2019) [arXiv:1806.00532 [astro-ph.IM]] [Search INSPIRE].
[39] T. Akutsu et al. [KAGRA Collaboration], Class. Quantum Grav. **36**, 165008 (2019) [arXiv:1901.03569 [astro-ph.IM]] [Search INSPIRE].