

Toward the detection of gravitational waves under non-Gaussian noises I. Locally optimal statistic

By Jun'ichi YOKOYAMA^{*1,*2,†}

(Communicated by Yoshihide KOZAI, M.J.A.)

Abstract: After reviewing the standard hypothesis test and the matched filter technique to identify gravitational waves under Gaussian noises, we introduce two methods to deal with non-Gaussian stationary noises. We formulate the likelihood ratio function under weakly non-Gaussian noises through the Edgeworth expansion and strongly non-Gaussian noises in terms of a new method we call Gaussian mapping where the observed marginal distribution and the two-body correlation function are fully taken into account. We then apply these two approaches to Student's t -distribution which has a larger tails than Gaussian. It is shown that while both methods work well in the case the non-Gaussianity is small, only the latter method works well for highly non-Gaussian case.

Keywords: Gravitational waves, data analysis, non-Gaussian noises, likelihood ratio, locally optimal statistic, Edgeworth expansion

1. Introduction

The large-scale cryogenic gravitational wave telescope (LCGT) now known as KAGRA started its construction in 2010 in deep underground at Kamioka mine. Recently its tunnel excavation, which started in early 2012,¹⁾ was successfully completed,²⁾ and the installation of a laser interferometer³⁾ has started toward the first direct detection of gravitational waves (GWs) in competition or cooperation with the advanced LIGO⁴⁾ and the advanced Virgo⁵⁾ detectors one century after Einstein proposed general relativity and predicted gravitational waves which propagate with the speed of light.

Among the four known elementary interactions in nature, gravity is by far the weakest force. Its high penetrating power would convey informations on deep inside celestial bodies and the very early Universe. On the other hand, this same property

also makes it very difficult to catch the signal. This is why no one has ever succeeded in direct detection of gravitational waves. Indeed it is a problem to extract a tiny signal out of much larger detector noise, and it is attempted by using a matched filter of expected signals. If detector noise is distributed Gaussian, we have a fairly straight forward technique as reviewed in the next section. However, the actual noises are known to be highly non-Gaussian. Therefore we must invent appropriate methods to deal with these non-Gaussianity so that we would not miss the signal.

In this article, first we present a mini review of signal detection under Gaussian noise, and then introduce two methods toward detection of gravitational waves under non-Gaussian noises to prepare for the forthcoming KAGRA. So far several papers have been written to deal with non-Gaussian noises for GW observations.^{6)–10)} However, they are mostly based on some specific non-Gaussian distributions such as double Gaussian, exponential,⁶⁾ χ^2 , and Student's t -distributions.⁸⁾ Here we try to be as model independent as possible, since we do not know the actual noise distribution a priori.

The rest of the paper is organized as follows. After a short review on signal detection under Gaussian noise in §2, in §3 we incorporate small deviation from Gaussian distribution in terms of the

^{*1} Research Center for the Early Universe (RESCEU), Graduate School of Science, The University of Tokyo, Tokyo, Japan.

^{*2} Kavli Institute for the Physics and Mathematics of the Universe (Kavli IPMU), TODIAS, WPI, The University of Tokyo, Kashiwa, Chiba, Japan.

[†] Correspondence should be addressed: J. Yokoyama, Research Center for the Early Universe (RESCEU), Graduate School of Science, The University of Tokyo, 113-0033, Japan (e-mail: yokoyama@resceu.s.u-tokyo.ac.jp).

Edgeworth expansion and calculate the likelihood ratio with it. In §4 the case with strong non-Gaussian noise is handled by a new method called Gaussian mapping, and the likelihood ratio test is formulated with arbitrary non-Gaussian marginal distribution. Then in §5 we apply the results of §§3 and 4 to a non-Gaussian distribution called Student's t -distribution which is a symmetric distribution with a larger tail than the corresponding Gaussian distribution. Finally §6 is devoted to conclusion.

2. Signal detection

Here we first present a minimal review on signal detection and describe optimal statistics and filter under Gaussian noise.*

2.1. Hypothesis testing. We first consider a hypothesis test using a time sequence of a detector output $x(t)$ and a GW signal (if any) $h(t)$. There are two distinct cases with and without nonvanishing signals $h(t)$. The former is expressed as H_1 or simply by 1 when we measure the sum $x(t) = n(t) + h(t)$ as the output, and the latter by H_0 or 0 measuring only noises $x(t) = n(t)$.

There are two kinds of errors associated with hypothesis testing. One is the false alarm (FA) claiming detection without actual signals, whose probability is given in terms of the conditional probability $P_{FA} = P(1|0)$. The other is false dismissal (FD) missing detection even if signal is there, given by the probability $P_{FD} = P(0|1)$. We wish to maximize the detection probability under a fixed FA probability or significance level.

An extremely useful theorem (although called lemma) is Neyman-Pearson's Lemma.¹⁵⁾ Using the likelihood ratio $\Lambda(x)$,

$$\Lambda(x) \equiv \frac{P(1|0)}{P(0|1)}, \quad [1]$$

we define a test such that

$$\begin{aligned} &\text{reject null hypothesis } H_0 \text{ if } \Lambda(x) > k \\ &\text{adopt null hypothesis } H_0 \text{ if } \Lambda(x) \leq k \end{aligned}$$

where k satisfies

$$P(\Lambda(x) > k|0) = \alpha.$$

Then the lemma states that this is the most powerful test at the significance level α . In other words, this test gives the largest detection rate under a fixed false alarm rate α . If there is a free fitting parameter, one

must first fix it by maximizing $\Lambda(x)$ before applying the test.

This lemma shows the importance of the likelihood ratio or the noise distribution function.

2.2. Signal detection under Gaussian noise.

Here we consider the case with random Gaussian noise. Discretizing the time sequence as $x(t_i) \equiv x_i$ the likelihood ratio reads

$$\begin{aligned} \Lambda(x) &= \exp \left[-\frac{1}{2} (x_i - h_i) (K^{-1})_{ij} (x_j - h_j) \right. \\ &\quad \left. + \frac{1}{2} x_i (K^{-1})_{ij} x_j \right] \\ &= \exp \left(q_i x_i - \frac{1}{2} h_i q_i \right) \end{aligned} \quad [2]$$

where $K_{ij} = \langle n_i n_j \rangle$ is the noise covariance matrix or the two-time correlation function, and $q_i \equiv (K^{-1})_{ij} h_j$ or $h_i = K_{ij} q_j$. Here and throughout, summation over repeated indices in the same term is assumed. In the continuum limit, we find

$$\langle n(t) n(t') \rangle \equiv K(t, t') \text{ and } h(t) = \int K(t, t') q(t') dt'. \quad [3]$$

Thus the log-likelihood ratio

$$\ln \Lambda(x) = \int q(t) x(t) dt - \frac{1}{2} \int q(t) h(t) dt \quad [4]$$

can be maximized if a linear correlator

$$G \equiv \int q(t) x(t) dt \quad [5]$$

is maximized. This is the linear matched filter for a known wave form $h(t)$.

Let us summarize some properties of this matched filter. The expectation value without any signal is simply equal to zero, $E_0\{G\} = 0$, whereas that with signal $h(t)$ reads

$$E_1\{G\} = \int q(t) h(t) dt. \quad [6]$$

The variance is given by

$$\text{Var}\{G\} = \int q(t) h(t) dt \quad [7]$$

irrespective of whether signal exists or not. We thus find the signal-to-noise (SN) ratio

$$\frac{S}{N} = \frac{E_1\{G\}}{\sqrt{\text{Var}\{G\}}} = \sqrt{\int q(t) h(t) dt}. \quad [8]$$

From this result we deduce that when the linear matched filter is maximized, the S/N ratio is also maximized for Gaussian noises.

* Standard references for this section include (Refs. 11, 12, 13, 14).

2.3. Frequency domain. It is often useful to work in the Fourier space with frequency f rather than dealing time sequence directly. Assuming stationary noise, its correlation function has time translational invariance, and we find

$$\begin{aligned}\langle n(t)n(t') \rangle &= K(t-t') \\ &= \frac{1}{2} \int_{-\infty}^{\infty} S_n(|f|) e^{-2\pi i f(t-t')} df, \quad [9]\end{aligned}$$

where $S_n(|f|) = 2S_n(f)$ is the one-sided power spectrum of noise. From

$$\begin{aligned}h(t) &= \int \tilde{h}(f) e^{-2\pi i f t} df \\ &= \int K(t-t') q(t') dt' \\ &= \frac{1}{2} \int S_n(|f|) \tilde{q}(f) e^{-2\pi i f(t-t')} df \quad [10]\end{aligned}$$

we find

$$\tilde{q}(f) = \frac{2\tilde{h}(f)}{S_n(|f|)}. \quad [11]$$

We therefore obtain

$$G = \int q(t)x(t)dt = \int \frac{2\tilde{h}(f)\tilde{x}^*(f)}{S_n(|f|)} df, \quad [12]$$

and the log-likelihood ratio

$$\begin{aligned}\ln \Lambda(x) &= \int \frac{2\tilde{h}(f)\tilde{x}^*(f)}{S_n(|f|)} df \\ &\quad - \frac{1}{2} \int \frac{2\tilde{h}(f)\tilde{h}^*(f)}{S_n(|f|)} df, \quad [13]\end{aligned}$$

for Gaussian noise.

2.4. Optimal filter. One can consider the question which filter $F(t)$ is optimal, or maximizes the S/N ratio. Defining

$$G_F \equiv \int F(t)x(t)dt = \int \tilde{F}(f)\tilde{x}^*(f)df,$$

we find the expectation value in the presence of a signal and the dispersion of G_F are respectively given by

$$\begin{aligned}E_1\{G_F\} &= \int \tilde{F}(f)\tilde{h}^*(f)df, \\ \text{Var}\{G_F\} &= \frac{1}{2} \int S_n(|f|)\tilde{F}(f)\tilde{F}^*(f)df. \quad [14]\end{aligned}$$

If we define an inner product by

$$[\tilde{a}(f), \tilde{b}(f)] \equiv \frac{1}{2} \int S_n(|f|)\tilde{a}(f)\tilde{b}^*(f)df,$$

[14] is expressed as

$$\begin{aligned}E_1\{G_F\} &= \left[\tilde{F}(f), \frac{2\tilde{h}(f)}{S_n(|f|)} \right], \\ \text{Var}\{G_F\} &= [\tilde{F}(f), \tilde{F}(f)],\end{aligned}$$

so that the S/N ratio is expressed as

$$\left(\frac{S}{N}\right)^2 = \frac{\left[\tilde{F}(f), \frac{2\tilde{h}(f)}{S_n(|f|)} \right]^2}{[\tilde{F}(f), \tilde{F}(f)]}. \quad [15]$$

This shows that the S/N ratio is maximized when $\tilde{F}(f) = 2\tilde{h}(f)/S_n(|f|)$. That is, the linear matched filter [12] used in the Gaussian likelihood function [13] maximizes the S/N ratio.

Note, however, that the above property holds if and only if noise distribution is Gaussian. For generic non-Gaussian distributions, there is no reason that the linear matched filter [12] is optimal. We should instead consider some non-linear filters in general. In the next two sections, instead, we consider the likelihood ratio test under non-Gaussian noises directly.

2.5. Locally optimal statistic. Before proceeding, however, we define a locally optimal statistic.¹⁶⁾ In the actual detection of GWs, the wave form has a number of undetermined parameters. As mentioned above, we maximize the likelihood function over these parameters. Among these undetermined parameter is an overall amplitude of GWs, ε , which depends on the distance to the source. Let us consider the case it is the only remaining parameter, quantifying N discrete sequence of the data as $x_i = n_i + \varepsilon \hat{h}_i$ with $\sum_{i=1}^N |\hat{h}_i|^2 = 1$.

Then the likelihood ratio depends on ε as $\Lambda(x; \varepsilon) = P(x|\varepsilon)/P(x|0)$. If the amplitude was large enough we could detect GW without any sophisticated statistical treatments, so what matters as a likelihood test is the case ε is very small. We may therefore expand the likelihood ratio with respect to ε as

$$\Lambda(x; \varepsilon) = 1 + \varepsilon \Lambda_1(x) + \frac{\varepsilon^2}{2} \Lambda_2(x) + \dots$$

The locally optimal statistic is defined by the first-order coefficient

$$\begin{aligned}\Lambda_1(x) &= \frac{1}{P(x|0)} \frac{d}{d\varepsilon} P(x|\varepsilon) \Big|_{\varepsilon=0} \\ &= \frac{d}{d\varepsilon} \ln \Lambda(x; \varepsilon) \Big|_{\varepsilon=0}, \quad [16]\end{aligned}$$

and it controls the entire likelihood ratio in case GW amplitude is small.

3. Signal detection under weakly non-Gaussian noise distribution

Since Gaussian noise is fully characterized by the covariance matrix or the two-body correlation function, any nonvanishing higher-order cumulants or reduced correlation functions are signatures of non-Gaussianity of the probability distribution function (PDF), $P(x)$. We incorporate effects of these nonvanishing higher-order cumulants to PDF and likelihood function to find the locally optimal statistics. The Edgeworth expansion¹⁷⁾ provides such a framework to incorporate them around the normal distribution, $\varphi(y) \equiv e^{-y^2/2}/\sqrt{2\pi}$, in relation with the central limit theorem. Here and hereafter, the variable y (as well as y_i or y_p introduced later) denotes a normalized quantity with a vanishing mean and a unit variance such as $y \equiv x/\sigma_x$ with $\sigma_x \equiv \sqrt{\langle x^2 \rangle}$.

First we expand $P(y)$ in terms of $\varphi(y)$ and its derivative as

$$\begin{aligned} P(y) &= \sum_{r=0}^{\infty} \frac{c_r}{r!} \varphi^{(r)}(y) \\ &= \sum_{r=0}^{\infty} \frac{c_r}{r!} (-)^r H_r(y) \varphi(y) \end{aligned} \quad [17]$$

where the second equality follows from the definition of the Hermite polynomial. From its orthonormality, the coefficients are given by

$$c_r = (-)^r \int_{-\infty}^{\infty} H_r(y) P(y) dy. \quad [18]$$

Using

$$\begin{aligned} H_0(y) &= 1, & H_1(y) &= y, & H_2(y) &= y^2 - 1, \\ H_3(y) &= y^3 - 3y, & H_4(y) &= y^4 - 6y^2 + 3, \dots \end{aligned} \quad [19]$$

we find

$$\begin{aligned} c_0 &= 1, & c_1 &= 0, & c_2 &= 0, \\ c_3 &= -\langle y^3 \rangle, & c_4 &= \langle y^4 \rangle - 3, \dots \end{aligned} \quad [20]$$

Now let us consider a random variable ξ which is a sum of \tilde{N} statistically independent variables with the same mean $\langle \xi_j \rangle = m_1$ and the variance $\langle (\xi_j - m_1)^2 \rangle = \sigma_1^2$ ($j = 1, 2, \dots, \tilde{N}$) as

$$\xi = \xi_1 + \xi_2 + \dots + \xi_{\tilde{N}}.$$

The central limit theorem asserts that the PDF, $P(y)$, of a variable $y = (\xi - m)/\sigma$, where $m = \langle \xi \rangle$ and $\sigma^2 = \langle (\xi - m)^2 \rangle = \tilde{N}\sigma_1^2$ are the mean and the variance of ξ , approaches the normal distribution as \tilde{N} increases.

We can relate the characteristic function for the PDF $P(y)$, $\Psi(z)$, with that of one component PDF, $P_1(y_i)$ with $y_i \equiv (\xi_i - m_1)/\sigma_1$ as

$$\begin{aligned} \Psi(z) &\equiv \int_{-\infty}^{\infty} e^{izy} P(y) dy = \langle e^{iz \frac{\xi - m}{\sigma}} \rangle \\ &= \left\langle \exp \left[iz \sum_{j=1}^{\tilde{N}} \frac{\xi_j - m_1}{\sqrt{\tilde{N}}\sigma_1} \right] \right\rangle \\ &= \left[\Psi_1 \left(\frac{z}{\sqrt{\tilde{N}}} \right) \right]^{\tilde{N}}, \end{aligned} \quad [21]$$

where $\Psi_1(z)$ is the characteristic function of $P_1(y_i)$. Now we define the following expansion

$$e^{\frac{z}{\sqrt{\tilde{N}}}} \Psi(z) = \int_{-\infty}^{\infty} e^{\frac{z}{\sqrt{\tilde{N}}} + izy} P(y) dy \equiv \sum_{m=0}^{\infty} \frac{\tilde{c}_m}{m!} (-iz)^m \quad [22]$$

and compare with [17]. Then using the identity,

$$\int e^{izy} \varphi^{(m)}(y) dy = (-iz)^m e^{-\frac{z^2}{2}}, \quad [23]$$

as well as the following relation between the characteristic function and the generating function, Φ_1 , of y_1 's cumulant λ_m ,

$$\Psi_1(z) = e^{\Phi_1(z)}, \quad \Phi_1(z) = \sum_{m=0}^{\infty} \frac{\lambda_m}{m!} z^m, \quad [24]$$

we find

$$\begin{aligned} e^{\frac{z}{\sqrt{\tilde{N}}}} \Psi(z) &= \sum_{m=0}^{\infty} \frac{c_m}{m!} (-iz)^m \\ &= \exp \left[\tilde{N} \sum_{j=3}^{\infty} \frac{\lambda_j}{j!} \left(\frac{iz}{\sqrt{\tilde{N}}} \right)^j \right] \\ &= \sum_{\ell=0}^{\infty} \frac{\tilde{N}^{\ell}}{\ell!} \left[\sum_{j=3}^{\infty} \frac{\lambda_j}{j!} \left(\frac{iz}{\sqrt{\tilde{N}}} \right)^j \right]^{\ell}. \end{aligned} \quad [25]$$

Then the j -th coefficient can be expressed by j -th and lower-order normalized cumulants and inverse powers of \tilde{N} ; for example,

$$\begin{aligned} c_3 &= -\frac{\lambda_3}{\sqrt{\tilde{N}}}, & c_4 &= \frac{\lambda_4}{\tilde{N}}, \\ c_5 &= -\frac{\lambda_5}{\tilde{N}^{3/2}}, & c_6 &= \frac{\lambda_6}{\tilde{N}^2} + 10 \frac{\lambda_3^2}{\tilde{N}}, \dots \end{aligned} \quad [26]$$

The Edgeworth expansion is obtained by rearranging these coefficients in powers of $\tilde{N}^{-1/2}$ as

$$\begin{aligned} P(y) &= \varphi(y) - \frac{\gamma_3}{3!} \varphi^{(3)}(y) + \frac{\gamma_4}{4!} \varphi^{(4)}(y) \\ &\quad + \frac{10\gamma_3^2}{6!} \varphi^{(6)}(y) + \dots, \end{aligned} \quad [27]$$

or equivalently,

$$P(y) = \left[1 + \frac{\gamma_3}{3!} H_3(y) + \frac{\gamma_4}{4!} H_4(y) + \frac{10\gamma_3^2}{6!} H_6(y) + \dots \right] \varphi(y), \quad [28]$$

where $\gamma_3 \equiv \lambda_3/\sqrt{\tilde{N}}$ and $\gamma_4 \equiv \lambda_4/\tilde{N}$.

This is the way we can incorporate higher-order cumulants around the otherwise Gaussian distribution. We note that although the derivation here is based on the approach to the Gaussian distribution on the basis of the central limit theorem, the same type Edgeworth expansion can also be found in the context of nonlinear evolution of density fluctuations starting from random Gaussian linear fluctuations.¹⁸⁾ In our problem, we of course put $y = x/\sigma$ where x is the detector output. But we may have a different control parameter than \tilde{N} which quantifies deviation from Gaussian depending on the nature of the underlying distribution.

Extention to a multivariate case around Gaussian PDF of N discretized time sequence of noises n_i is possible starting with the original noise PDF in the Gaussian limit,

$$P[\{n_i\}] d^N n = \frac{1}{\sqrt{(2\pi)^N \|K\|}} \exp \left[-\frac{1}{2} \sum_{j,\ell=1}^N n_j (K^{-1})_{j\ell} n_\ell \right] d^N n, \quad [29]$$

which can be diagonalized by a real unitary matrix U as $n_j = U_{j\ell} \psi_\ell$

$$\begin{aligned} n_j (K^{-1})_{j\ell} n_\ell &= \psi_j (U^\dagger K^{-1} U)_{j\ell} \psi_\ell \\ &= \sum_{p=1}^N \Lambda_p \psi_p^2 \equiv \sum_{p=1}^N y_p^2, \\ \Lambda_p &= \frac{1}{\sigma_p^2}, \quad y_p = \frac{\psi_p}{\sigma_p}, \end{aligned} \quad [30]$$

where Λ_p ($p = 1 - N$) are eigenvalues of the matrix K^{-1} . For each y_p we can apply the Edgeworth expansion independently, to yield

$$P_{EW}[\{y_i\}] = \prod_{p=1}^N \left[1 + \frac{\gamma_3^{(p)}}{3!} H_3(y_p) + \frac{\gamma_4^{(p)}}{4!} H_4(y_p) + \frac{10\gamma_3^{(p)2}}{6!} H_6(y_p) + \dots \right] \varphi(y_p) \quad [31]$$

The log likelihood function ratio reads

$$\begin{aligned} \ln \Lambda_{EW} &= (\text{Gaussian part}) + \sum_{p=1}^N \ln \left[1 + \frac{\gamma_3^{(p)}}{3!} H_3(\hat{y}_p) + \frac{\gamma_4^{(p)}}{4!} H_4(\hat{y}_p) + \frac{10\gamma_3^{(p)2}}{6!} H_6(\hat{y}_p) + \dots \right] \\ &\quad - \sum_{p=1}^N \ln \left[1 + \frac{\gamma_3^{(p)}}{3!} H_3(y_p) + \frac{\gamma_4^{(p)}}{4!} H_4(y_p) + \frac{10\gamma_3^{(p)2}}{6!} H_6(y_p) + \dots \right] \end{aligned} \quad [32]$$

where

$$y_p = \frac{\psi_p}{\sigma_p} = \frac{1}{\sigma_p} U_{pj}^\dagger n_j, \quad \hat{y}_p = \frac{1}{\sigma_p} U_{pj}^\dagger (n_j - \epsilon \hat{h}_j)$$

$$\begin{aligned} \frac{d \ln \Lambda_{EW}}{d\epsilon} \Big|_{\epsilon=0} &= (\text{Gaussian part}) \\ &\quad - \sum_{p=1}^N \frac{\frac{\gamma_3^{(p)}}{3!} H'_3(y_p) + \frac{\gamma_4^{(p)}}{4!} H'_4(y_p) + \dots}{1 + \frac{\gamma_3^{(p)}}{3!} H_3(y_p) + \frac{\gamma_4^{(p)}}{4!} H_4(y_p) + \dots} U_{pj}^\dagger \frac{\hat{h}_j}{\sigma_p} \end{aligned} \quad [33]$$

If we expand logarithm in the right hand side of [32], which is valid if deviation from Gaussian is small, we find

$$\begin{aligned} \ln \Lambda_{EW} &\simeq (\text{Gaussian part}) + \sum_{p=1}^N \frac{\gamma_3^{(p)}}{3!} [H_3(\hat{y}_p) - H_3(y_p)] \\ &\quad + \sum_{p=1}^N \frac{\gamma_4^{(p)}}{4!} [H_4(\hat{y}_p) - H_4(y_p)] + \dots \end{aligned}$$

From this expression, the locally optimal statistic reads

$$\begin{aligned} \frac{d \ln \Lambda_{EW}}{d\epsilon} \Big|_{\epsilon=0} &\simeq (\text{Gaussian part}) \\ &\quad - \sum_{p=1}^N \left[\frac{\gamma_3^{(p)}}{2} (y_p^2 - 1) + \frac{\gamma_4^{(p)}}{6} (y_p^3 - 3y_p) \right] U_{pj}^\dagger \frac{\hat{h}_j}{\sigma_p} + \dots \end{aligned} \quad [34]$$

For the case of stationary noise, we can obtain more explicit expression by virtue of the discrete Fourier transformation. Suppose that the data is sampled with a time interval Δ from $t = 0$ to $T_{obs} \equiv N\Delta$, namely, at $t_j = j\Delta$ ($j = 1, 2, \dots, N$). Then the discrete Fourier transform of $n(t_j)$ is given as

$$\tilde{n}(f_m) = \Delta \sum_{j=1}^N n(t_j) e^{2\pi i \frac{jm}{N}} \quad [35]$$

at $f_m \equiv \frac{m}{N\Delta}$ with $m = -\frac{N}{2}, -\frac{N}{2}+1, \dots, 0, 1, \dots, \frac{N}{2}$, assuming N is an even number. Its inverse reads

$$n(t_j) = \frac{1}{N\Delta} \sum_{m=1}^N \tilde{n}(f_m) e^{-2\pi i \frac{jm}{N}}. \quad [36]$$

In the stationary Gaussian PDF

$$\begin{aligned} P[\{n_i\}] d^N n &= \frac{1}{\sqrt{(2\pi)^N \|K\|}} \\ &\times \exp \left[-\frac{1}{2} \sum_{j,\ell=1}^N n(t_j) K^{-1}(t_j - t_\ell) n(t_\ell) \Delta^2 \right] d^N n, \end{aligned} \quad [37]$$

we can expand the inverse covariance function as

$$K^{-1}(t_j - t_\ell) = \frac{1}{N\Delta} \sum_{p=1}^N \widetilde{K^{-1}}(f_p) e^{-2\pi i(j-\ell)\frac{p}{N}} \quad [38]$$

so that we find

$$\begin{aligned} &\sum_{j,\ell=1}^N n(t_j) K^{-1}(t_j - t_\ell) n(t_\ell) \Delta^2 \\ &= \frac{1}{N^3 \Delta^3} \sum_{m,p,q=1}^N \tilde{n}(f_m) \widetilde{K^{-1}}(f_p) \tilde{n}(f_q) \Delta^2 N \delta_{m+p,N} N \delta_{q-p,N} \\ &= \frac{1}{N\Delta} \sum_{p=1}^N \tilde{n}(f_{N-p}) \widetilde{K^{-1}}(f_p) \tilde{n}(f_p) \\ &= \frac{1}{N\Delta} \sum_{p=1}^N \widetilde{K^{-1}}(f_p) |\tilde{n}(f_p)|^2 \end{aligned} \quad [39]$$

where we have used an identity $\tilde{n}(f_p) = \tilde{n}^*(f_{N-p})$.

Using the discretized version of eq. [9], we find

$$\begin{aligned} \widetilde{K^{-1}}(f_p) &= \frac{2}{S_n(f_p)}, \\ \text{and } \langle |\tilde{n}(f_p)|^2 \rangle &= \frac{1}{2} S_n(f_p) T_{obs} \equiv 2\tilde{\sigma}_p^2. \end{aligned} \quad [40]$$

It is simpler to consider two real valued quantities, $\tilde{n}_R(f_p) \equiv \text{Re } \tilde{n}(f_p)$ and $\tilde{n}_I(f_p) \equiv \text{Im } \tilde{n}(f_p)$ than dealing with a complex variable $\tilde{n}(f_p)$ itself. Under the distribution [37], these variables satisfy $\langle \tilde{n}_R^2(f_p) \rangle = \langle \tilde{n}_I^2(f_p) \rangle = S_n(f_p) T_{obs}/4$ and $\langle \tilde{n}_R(f_p) \tilde{n}_I(f_p) \rangle = 0$, and they do not have any correlation with modes of different frequencies.

Hence we can identify

$$y_{p1} \equiv \frac{\tilde{n}_R(f_p)}{\tilde{\sigma}_p} \quad \text{and} \quad y_{p2} \equiv \frac{\tilde{n}_I(f_p)}{\tilde{\sigma}_p} \quad [41]$$

in [30]. Then the locally optimal statistic corresponding to [34] reads

$$\begin{aligned} &\frac{d \ln \Lambda_{EW}}{d\epsilon} \Big|_{\epsilon=0} \\ &\simeq (\text{Gaussian part}) - \sum_{i=1}^2 \sum_{p_i=1}^N \left[\frac{\gamma_3^{(p_i)}}{2} (y_{p_i}^2 - 1) \right. \\ &\quad \left. + \frac{\gamma_4^{(p_i)}}{6} (y_{p_i}^3 - 3y_{p_i}) + \dots \right] \frac{\tilde{h}_i(f_{p_i})}{\tilde{\sigma}_{p_i}}, \end{aligned} \quad [42]$$

with $\tilde{h}_1(f_{p_i}) \equiv \text{Re } \tilde{h}(f_{p_i})$ and $\tilde{h}_2(f_{p_i}) \equiv \text{Im } \tilde{h}(f_{p_i})$ where $\tilde{h}(f_{p_i})$ is the Fourier transform of $\hat{h}(t_j)$.

4. Gaussian mapping

The Edgeworth expansion discussed in the previous section works if and only if deviation from Gaussian is small. In order to deal with more realistic cases with larger deviation from Gaussian, we introduce the following new method which we call Gaussian mapping. In this approach the observed one-point PDF or the marginal distribution can be an arbitrary non-Gaussian distribution and also two body correlation function can be fully taken into account. Since it is a formidable task to determine higher order correlation functions of noises observationally, this method makes use of and reproduces as much observational information as possible.

We start with a multivariate random Gaussian

$$\begin{aligned} &P_\phi[\{\phi_i\}] d^N \phi \\ &= \frac{1}{\sqrt{(2\pi)^N \|\zeta\|}} \exp \left[-\frac{1}{2} \phi_j (\zeta^{-1})_{j\ell} \phi_\ell \right] d^N \phi \end{aligned} \quad [43]$$

with

$$\zeta_{j\ell} = \langle \phi_j \phi_\ell \rangle, \quad \zeta_{jj} = 1, \quad |\zeta_{j\ell}| \leq 1,$$

where $\|\zeta\|$ denotes the determinant of the covariance matrix ζ_{ij} . Suppose that the noise $n(t_i)$ is a function of ϕ_i , $n(t_i) = Q[\phi_i]$ at each time and that there exists an inverse function of Q , Q^{-1} so that

$$\begin{aligned} \phi_i &= Q^{-1}[n(t_i)] \equiv g[n(t_i)] = g(n_i), \\ d\phi_i &= g'(n_i) dn_i. \end{aligned} \quad [44]$$

Then the multivariate PDF for n_i reads

$$\begin{aligned} &P[\{n_i\}] d^N n \\ &= \frac{1}{\sqrt{(2\pi)^N \|\zeta\|}} \exp \left[-\frac{1}{2} g(n_j) (\zeta^{-1})_{j\ell} g(n_\ell) \right] \prod_{p=1}^N g'(n_p) dn_p \end{aligned} \quad [45]$$

The one-point PDF reads

$$\begin{aligned} P(n_j)dn_j &= P_\phi[\phi_j = Q^{-1}(n_j)]g'(n_j)dn_j \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}g^2(n_j)} g'(n_j)dn_j. \end{aligned} \quad [46]$$

This is to be determined by observations of the “training” samples. The cumulative PDF

$$\hat{P}(n) \equiv \int_0^n P(n')dn'$$

is given by

$$\begin{aligned} \hat{P}(n) &= \frac{1}{\sqrt{2\pi}} \int_0^{g(n)} e^{-\frac{1}{2}y^2} dy = \frac{1}{\sqrt{\pi}} \text{Erf}\left(\frac{g(n)}{\sqrt{2}}\right), \\ \text{Erf}x &\equiv \int_0^x e^{-t^2} dt. \end{aligned} \quad [47]$$

We therefore find

$$\begin{aligned} g(n) &= \sqrt{2}\text{Erf}^{-1}(\sqrt{\pi}\hat{P}(n)) \\ &= \sqrt{2}\text{Erf}^{-1}\left(\sqrt{\pi} \int_0^n P(n')dn'\right). \end{aligned} \quad [48]$$

For an arbitrary well behaved one point PDF, $P(n)$, we can find a function $g(n)$ as above.

Next we incorporate two-point correlation function, which is important when noises cannot be “whitened” in a pre-process of data analysis pipeline with a relatively long correlation time.

$$\begin{aligned} \langle n(t_m)n(t_n) \rangle &= \langle Q[\phi_m]Q[\phi_n] \rangle_{\phi-\text{Gaussian}} \\ &= \int d^N\phi \frac{Q(\phi_m)Q(\phi_n)}{\sqrt{(2\pi)^N \|\zeta\|}} \exp\left[-\frac{1}{2}\phi_j(\zeta^{-1})_{j\ell}\phi_\ell\right]. \end{aligned} \quad [49]$$

Using Fourier transform

$$Q(\phi_j) = \int dk_j e^{-2\pi i k_j \phi_j} \tilde{Q}(k_j),$$

we find

$$\begin{aligned} \langle n(t_m)n(t_n) \rangle &= \int dk_m dk_n \frac{\tilde{Q}(k_m)\tilde{Q}^*(k_n)}{\sqrt{(2\pi)^N \|\zeta\|}} \\ &\quad \times \exp\left[-2\pi i(k_m\phi_m - k_n\phi_n) - \frac{1}{2}\phi_j(\zeta^{-1})_{j\ell}\phi_\ell\right] d\phi_m d\phi_n \\ &= \int dk_m dk_n \tilde{Q}(k_m)\tilde{Q}^*(k_n) \\ &\quad \times \exp[-2\pi^2(k_m^2 - 2k_mk_n\zeta_{mn} + k_n^2)] \\ &= \int \frac{1}{2}\tilde{Q}\left(\frac{u+v}{2}\right)\tilde{Q}\left(\frac{u-v}{2}\right) \\ &\quad \times \exp[-\pi^2(1 - \zeta_{mn})u^2 - \pi^2(1 + \zeta_{mn})v^2] du dv \end{aligned} \quad [50]$$

In case n is pure Gaussian we find $n = Q(\phi) = \sigma\phi$ with σ being the dispersion of the noise. Then the

Fourier transform of $Q(\phi)$ is given by

$$\tilde{Q}(k) = \frac{\sigma}{2\pi i} \delta'(k), \quad [52]$$

which reproduces $\langle n(t_m)n(t_n) \rangle = \sigma^2\zeta_{mn}$ from [50] as it should be.

If, on the other hand, $\tilde{Q}(k)$ is a smooth function around $k=0$, we can estimate [51] in terms of a saddle point approximation to yield

$$\begin{aligned} \langle n(t_m)n(t_n) \rangle &\equiv K(t_m, t_n) \\ &\simeq \frac{|\tilde{Q}(0)|^2}{2\pi^3} \frac{1}{\sqrt{1 - \zeta_{mn}^2}}. \end{aligned} \quad [53]$$

So if we obtain $K(t_m, t_n) = K_{mn}$ observationally, we can find ζ_{mn} as

$$\zeta_{mn} \simeq \left[1 - \left(\frac{|\tilde{Q}(0)|^2}{2\pi^3 K_{mn}}\right)^2\right]^{1/2}. \quad [54]$$

In the actual application, we should evaluate [51] numerically to relate $\langle n(t_m)n(t_n) \rangle$ with ζ_{mn} . Note that if noise is stationary, ζ_{mn} is also stationary. Once these correspondences are achieved, one can calculate the log likelihood function as

$$\begin{aligned} \ln \Lambda_{GM} &= \sum_{j,\ell=1}^N \left[-\frac{1}{2} g(x_j - \epsilon\hat{h}_j)(\zeta^{-1})_{j\ell} g(x_\ell - \epsilon\hat{h}_\ell) \right. \\ &\quad \left. + \frac{1}{2} g(x_j)(\zeta^{-1})_{j\ell} g(x_\ell) \right] + \sum_{j=1}^N \ln \frac{g'(x_j - \epsilon\hat{h}_j)}{g'(x_j)}, \end{aligned} \quad [55]$$

and the locally optimal statistic reads

$$\begin{aligned} \frac{d \ln \Lambda_{GM}}{d\epsilon} \Big|_{\epsilon=0} &= \sum_{j,\ell=1}^N g'(x_j)(\zeta^{-1})_{j\ell} g(x_\ell) \hat{h}_j \\ &\quad - \sum_{j=1}^N \frac{g''(x_j)}{g'(x_j)} \hat{h}_j. \end{aligned} \quad [56]$$

So far we have developed a new formalism primarily for real time analysis dealing with a time sequence of data directly. This method, however, may be applied to the analysis in the frequency domain, too, if we assume one-point PDF of each normalized Fourier mode, $\tilde{n}(f)/\sqrt{\langle |\tilde{n}(f)|^2 \rangle}$, satisfies the same distribution function in analogy to [44]. In the next section we employ a specific non-Gaussian model to test the performance of the Edgeworth expansion and Gaussian mapping based on the above treatment.

5. Application to Student's t -distribution

Here we compare the results of the previous two sections using a specific non-Gaussian model, namely,

Student's t -distribution as an example. This distribution is an even function with a longer and larger tails than the Gaussian distribution with the same variance. Let us work in the Fourier space and consider a single frequency mode, $\tilde{x}(f)$, assuming that there is no correlation with the other frequencies. Let us further assume that the noise has random phase, so that $\text{Re } \tilde{n}(f) \equiv \tilde{n}_R$ and $\text{Im } \tilde{n}(f) \equiv \tilde{n}_I$ have no bilinear correlation, $\langle \tilde{n}_R(f) \tilde{n}_I(f) \rangle = 0$. Then these two variables satisfy a bivariate t -distribution.

$$P_2(\tilde{n}_R, \tilde{n}_I) d\tilde{n}_R d\tilde{n}_I = \frac{1}{2\pi} \frac{m}{(m-2)\tilde{\sigma}^2} \left[1 + \frac{\tilde{n}_R^2 + \tilde{n}_I^2}{(m-2)\tilde{\sigma}^2} \right]^{-\frac{m+2}{2}} d\tilde{n}_R d\tilde{n}_I, \quad [57]$$

where $\tilde{\sigma}^2 \equiv \langle \tilde{n}_R^2 \rangle = \langle \tilde{n}_I^2 \rangle = \langle |\tilde{n}(f)|^2 \rangle / 2$ and m , which we assume is larger than 4, denotes a parameter called the degree of freedom. This distribution approaches Gaussian as we take $m \rightarrow \infty$.

Let us consider the normalized variables $y_1 \equiv \tilde{n}_R/\tilde{\sigma}$ and $y_2 \equiv \tilde{n}_I/\tilde{\sigma}$ with the PDF

$$P_{2t}(y_1, y_2) dy_1 dy_2 = \frac{m}{2\pi(m-2)} \left(1 + \frac{y_1^2 + y_2^2}{m-2} \right)^{-\frac{m+2}{2}} dy_1 dy_2, \quad [58]$$

hereafter.

The locally optimal statistic calculated from this PDF is given by

$$\frac{d \ln \Lambda_{2t}(y_1, y_2)}{d\epsilon} \Big|_{\epsilon=0} = \frac{m+2}{m-2+y_1^2+y_2^2} \left(y_1 \frac{\tilde{h}_1}{\tilde{\sigma}} + y_2 \frac{\tilde{h}_2}{\tilde{\sigma}} \right), \quad [59]$$

again with $\tilde{h}_1 \equiv \text{Re } \tilde{h}(f)$ and $\tilde{h}_2 \equiv \text{Im } \tilde{h}(f)$.

The Edgeworth expansion of [58] can be manipulated straight forwardly thanks to the random phase nature which yields $\langle y_1 y_2 \rangle = 0$. The fact that odd-order cumulants all vanish makes the expression even simpler. In the t -distribution, the inverse of the degree of freedom, $1/m$, acts as the expansion parameter to quantify the approach to the Gaussian distribution in large m limit. Up to $\mathcal{O}(m^{-3/2})$ we find

$$\begin{aligned} P_{EW}(y_1, y_2) &= \left[1 + \frac{\kappa_4}{4!} H_4(y_1) + \dots \right] \\ &\times \left[1 + \frac{\kappa_4}{4!} H_4(y_2) + \dots \right] \varphi(y_1) \varphi(y_2) \\ &\cong \left[1 + \frac{\kappa_4}{4!} (y_1^4 - 6y_1^2 + y_2^4 - 3y_2^2 + 6) \right] \varphi(y_1) \varphi(y_2). \end{aligned} \quad [60]$$

Here κ_4 is the fourth cumulant of y_1 and y_2 calculated by their marginal distribution

$$\begin{aligned} P_{1t}(y) &= \int P_{2t}(y, y_2) dy_2 \\ &= \frac{1}{\sqrt{(m-2)\pi}} \frac{\Gamma\left(\frac{m+1}{2}\right)}{\Gamma\left(\frac{m}{2}\right)} \left(1 + \frac{y^2}{m-2} \right)^{-\frac{m+1}{2}}, \end{aligned} \quad [61]$$

that is, $\kappa_4 = \langle y_1^4 \rangle - 3 = \frac{6}{m-4}$.

As a result the locally optimal statistic corresponding to [33] is given by

$$\begin{aligned} \frac{d \ln \Lambda_{EW}}{d\epsilon} \Big|_{\epsilon=0} &= \left[y_1 - \frac{4y_1^3 - 12y_1}{4(m-4) + y_1^4 - 6y_1^2 + 3} \right] \frac{\tilde{h}_1}{\tilde{\sigma}} \\ &+ \left[y_2 - \frac{4y_2^3 - 12y_2}{4(m-4) + y_2^4 - 6y_2^2 + 3} \right] \frac{\tilde{h}_2}{\tilde{\sigma}} \end{aligned} \quad [62]$$

$$\simeq \left(1 + \frac{3-y_1^2}{m} \right) y_1 \frac{\tilde{h}_1}{\tilde{\sigma}} + \left(1 + \frac{3-y_2^2}{m} \right) y_2 \frac{\tilde{h}_2}{\tilde{\sigma}}, \quad [63]$$

the last expression being valid for $m \gg \max(4, y_1^4/4, y_2^4/4)$.

Now let us turn to the analysis in the Gaussian mapping method. In the two variable system at hand, thanks to the property $\langle y_1 y_2 \rangle = 0$ again, this method is equivalent with treating y_1 and y_2 as fully independent variables with no mutual correlations. Hence in this method the PDF is given by a product of the marginal distributions [61],

$$P_{GM}(y_1, y_2) = P_{1t}(y_1) P_{1t}(y_2). \quad [64]$$

One could in principle calculate $g(n)$ function introduced in the previous section to use the formula [56], but in practice the same result can be obtained by using [64] more directly in this case, to yield

$$\begin{aligned} \frac{d \ln \Lambda_{GM}}{d\epsilon} \Big|_{\epsilon=0} &= \frac{(m+1)y_1}{m-2+y_1^2} \frac{\tilde{h}_1}{\tilde{\sigma}} + \frac{(m+1)y_2}{m-2+y_2^2} \frac{\tilde{h}_2}{\tilde{\sigma}} \\ &\simeq \left(1 + \frac{3-y_1^2}{m} \right) y_1 \frac{\tilde{h}_1}{\tilde{\sigma}} + \left(1 + \frac{3-y_2^2}{m} \right) y_2 \frac{\tilde{h}_2}{\tilde{\sigma}}, \end{aligned} \quad [65]$$

the latter being valid for $m \gg \max(4, y_1^2, y_2^2)$. In this limit, the Edgeworth expansion and the Gaussian mapping yield the same result. However, when m , y_1 , and y_2 do not satisfy aforementioned inequalities, we must use [62] and [65] directly.

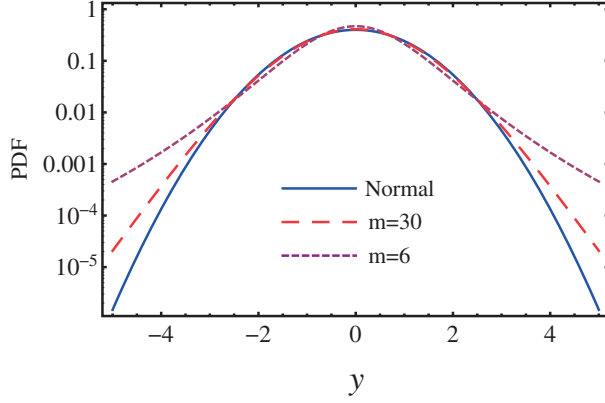


Fig. 1. Probability distribution functions of Normal distribution $\varphi(y)$ (solid line) and Student's t -distribution [61] with $m = 30$ (long dashed line) and 60 (short dashed line), both normalized to have a unit variance.

Let us study performances of the Edgeworth expansion [62] and the Gaussian mapping [65] comparing them with the result of the bivariate t -distribution [59]. First Fig. 1 shows single variate t -distribution with unit variance $\langle y^2 \rangle = 1$ for $m = 30$ and 6 together with the normal distribution. As is seen there t -distribution with smaller m has larger tail.

For simplicity of illustration, let us focus on the case $\tilde{h}_2 = 0$ and depict the dependence of the locally optimal statistic along y_1 direction. Figure 2 shows the locally optimal statistic, which is expressed in unit of $\tilde{h}_1/\tilde{\sigma}$, as a function of y_1 for the case $m = 30$. In this figure we have taken a somewhat atypical value of $y_2 = 2$ because for smaller $|y_2|$ the result of the Gaussian mapping is indistinguishable from that of the exact bivariate t -distribution [59]. As is seen there even the Edgeworth expansion works well for $|y_1| \lesssim 3$.

Figure 3, on the other hand, shows a highly non-Gaussian case with $m = 6$. There we have taken $y_2^2 = 1$ as a typical value of its variance. As is seen there, the agreement between the true bivariate t -distribution and the Gaussian mapping is striking. But the Edgeworth expansion works only in the small vicinity of $y_1 = 0$.

6. Conclusion

In the present paper, we have considered ways to handle with non-Gaussian natures of detector noises for the detection of gravitational waves with the forthcoming KAGRA or other large-scale laser interferometers. After reviewing the standard theory

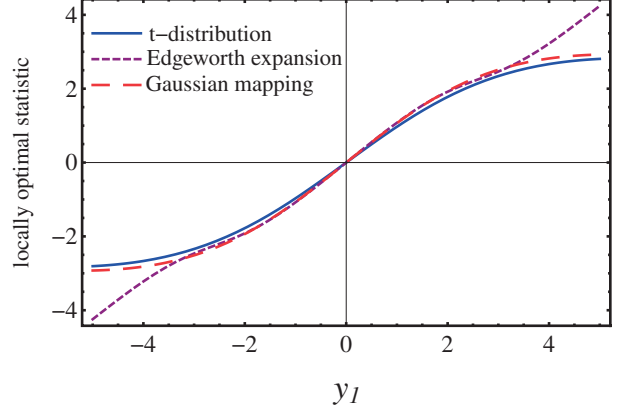


Fig. 2. Locally optimal statistics for Student's t -distribution (solid line), its Edgeworth expansion (short dashed line), and the result of Gaussian mapping (long dashed line) for the case $m = 30$ and $y_2^2 = 4$. The vertical axis is expressed in unit of $\tilde{h}_1/\tilde{\sigma}$.

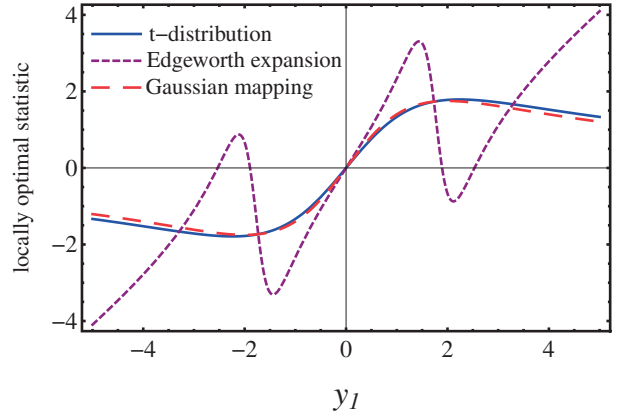


Fig. 3. Locally optimal statistics for Student's t -distribution (solid line), its Edgeworth expansion (short dashed line), and the result of Gaussian mapping (long dashed line) for the case $m = 6$ and $y_2^2 = 1$. The vertical axis is expressed in unit of $\tilde{h}_1/\tilde{\sigma}$.

of hypothesis testing and matched filter technique used in the conventional analysis of gravitational waves assuming Gaussian noises, we have presented two ways to calculate the likelihood ratio or the locally optimal statistic, which plays the central role in hypothesis testing, in the presence of stationary non-Gaussian noises.

One is the Edgeworth expansion which can incorporate weak deviation from Gaussian noises, and it has been used in various problems near the regime the central limit theorem works¹⁷⁾ as well as in the weakly nonlinear evolution of density fluctuations in the Universe.¹⁸⁾

The other is a new method which we call Gaussian mapping. In this formalism we first determine the single-time or the marginal probability distribution of noises using training samples as well as the two-time correlation function. Then we can formulate the likelihood ratio which fully incorporates these two data making use of mapping to Gaussian distribution. Since observational determination of higher-order correlation function is a formidable task, in this sense, this method makes the most use of the observationally available data of noise distribution. This method can also be applied to the analysis in frequency domain if each normalized Fourier mode satisfies the same mono-variate distribution function.

Applying these two methods for a specific non-Gaussian distribution called Student's t -distribution, which mimics actual noise distribution function in the sense that it is a symmetric distribution with a much heavier tails than Gaussian, we have shown that the Edgeworth expansion works well if and only if deviation from Gaussian is small but the formula based on the Gaussian mapping works very well even in highly non-Gaussian case.

These methods attempt to treat the stationary non-Gaussian parts of noises by frontal attack, but there is another approach which makes use of non-Gaussianity to separate signals from noises known as the independent component analysis.^{19)–21)} This method has not been used for the data analysis of gravitational waves so far since we need detailed information of envirometers such as seismograph which measure various independent noises. In a forthcoming paper²²⁾ we shall consider application of this method to the data analysis for the first time. We then plan to apply the analytic results of the present and the next papers to actual noises observed by existent detectors. By pursuing various analysis methods we shall prepare for the completion of KAGRA detector toward the first direct detection of gravitational waves.

Acknowledgements

The author is grateful to Yousuke Itoh, Nobuyuki Kanda, Masaki Ando, and Yuhei Miyamoto for useful communications. This work was partially supported by the Grant-in-Aid for Scientific Research on Innovative Areas No. 25103504.

References

- 1) Ohashi, M. (2012) Construction of KAGRA, the large-scale cryogenic gravitational wave telescope, Begins. AAPPs Bulletin **22** (2), 18–19.
- 2) Miyoki, S. and Uchiyama, T. (2014) Excavation of the tunnel for KAGRA gravitational wave telescope is now complete. AAPPs Bulletin **24** (4), 40–42.
- 3) Aso, Y. *et al.* [KAGRA collaboration] (2013) Interferometer design of the KAGRA gravitational wave detector. Phys. Rev. D Part. Fields **88**, 043007.
- 4) Harry, G.M. [LIGO Scientific Collaboration] (2010) Advanced LIGO: The next generation of gravitational wave detectors. Class. Quant. Grav. **27**, 084006.
- 5) Acernese, F. *et al.* [The Virgo Collaboration], arXiv:1408.3978 [gr-qc].
- 6) Allen, B., Creighton, J.D.E., Flanagan, E.E. and Romano, J.D. (2002) Robust statistics for deterministic and stochastic gravitational waves in non-Gaussian noise: Frequentist analysis. Phys. Rev. D Part. Fields **65**, 122002.
- 7) Allen, B., Creighton, J.D.E., Flanagan, E.E. and Romano, J.D. (2002) Robust statistics for deterministic and stochastic gravitational waves in non-Gaussian noise. II. Bayesian analysis. Phys. Rev. D Part. Fields **67**, 122002.
- 8) Röver, C. (2011) Student- t based filter for robust signal detection. Phys. Rev. D Part. Fields **84**, 122004.
- 9) Biswas, R. *et al.* (2012) Likelihood-ratio ranking of gravitational-wave candidates in a non-Gaussian background. Phys. Rev. D Part. Fields **85**, 122008.
- 10) Biswas, R. *et al.* (2012) Detecting transient gravitational waves in non-Gaussian noise with partially redundant analysis methods. Phys. Rev. D Part. Fields **85**, 122009.
- 11) Wainstein, L.A. and Zubakov, V.D. (1965) Extraction of Signals from Noise. Dover Publications, New York, USA.
- 12) Maggiore, M. (2008) Gravitational Waves Volume 1 Theory and Experiments. Oxford University Press, Oxford, UK.
- 13) Jaranowski, P. and Królak, A. (2009) Analysis of Gravitational-Wave Data. Cambridge University Press, Cambridge, UK.
- 14) Creighton, J.D.E. and Anderson, W.G. (2011) Gravitational-Wave Physics and Astronomy An Introduction to Theory, Experiment, and Data Analysis. Wiley-VCH Verlag, Weinheim, Germany.
- 15) Neyman, J. and Pearson, E.S. (1933) On the problem of the most efficient tests of statistical hypotheses. Philos. Trans. R. Soc. Lond. A **231**, 289–337.
- 16) See, *e.g.*, Kassam, S.A. (1988) Signal Detection in Non-Gaussian Noise. Springer-Verlag, Berlin, Germany.
- 17) Cramér, H. (1946) Mathematical Methods of Statistics (Princeton: Princeton University Press).
- 18) Juszkwicz, R., Weinberg, D.H., Amsterdamski, P., Chodorowski, M. and Bouchet, F. (1995) Weakly nonlinear Gaussian fluctuations and the Edgeworth expansion. Astrophys. J. **442**, 39–56.
- 19) Jutten, C. and Herault, J. (1991) Blind separation of sources, Part I: An adaptive algorithm based on

- neuromimetic architecture. *Signal Process.* **24**, 1–10.
- 20) Common, P. (1994) Independent component analysis, a new concept? *Signal Process.* **36**, 287–314.
- 21) Amari, S. and Cardoso, J.-F. (1997) Blind source separation — semiparametric statistical approach. *IEEE Trans. Signal Process.* **45**, 2692–2700.
- 22) Yokoyama, J. Toward the detection of gravitational waves under non-Gaussian noises II. Independent component analysis (to be published).

(Received July 30, 2014; accepted Oct. 25, 2014)