# Non-Gaussianity Detection of EEG Signals Based on a Multivariate Scale Mixture Model for Diagnosis of Epileptic Seizures

Akira Furui*, *Member, IEEE,* Ryota Onishi, Akihito Takeuchi, Tomoyuki Akiyama*, and Toshio Tsuji*, *Member, IEEE*

*Abstract*—*Objective:* The detection of epileptic seizures from scalp electroencephalogram (EEG) signals can facilitate early diagnosis and treatment. Previous studies suggested that the Gaussianity of EEG distributions changes depending on the presence or absence of seizures; however, there is no general EEG signal model that can explain such changes in distributions within a unified scheme. *Methods:* This paper describes the formulation of a stochastic EEG model based on a multivariate scale mixture distribution; the model can represent changes in non-Gaussianity caused by stochastic fluctuations in EEG. In addition, we propose a novel EEG analysis method by combining the model with a filter bank and introduce an index representing the non-Gaussianity latent in each EEG frequency band. *Results:* We applied the proposed method to multichannel EEG data from twenty patients with focal epilepsy. The results showed significant increases the proposed index during epileptic seizures, particularly in the high-frequency band. The index calculated in the high-frequency band allowed highly accurate classification of seizure segments and non-seizure segments (area under the curve (AUC) = 0.881) using only a simple threshold, and significantly outperformed the conventional amplitude-based index ($p < 0.001$). *Conclusion:* This study proposed a multivariate scale mixture distribution-based stochastic EEG model capable of representing non-Gaussianity associated with epileptic seizures. In the high-frequency band, the proposed index shows the possibility of classification of epileptic seizures. *Significance:* The proposed model can be used in various fields of EEG, and helps to characterize stochastic fluctuations of multichannel EEG signals.

*Index Terms*—Electroencephalogram (EEG), stochastic model, stochastic fluctuation, multivariate scale mixture model, epileptic seizure, non-Gaussianity.

## I. INTRODUCTION

**E**PILEPSY is a chronic neurological disorder that is characterized by recurrent and unprovoked epileptic seizures caused by abnormal electric activities of brain neurons. It is reported that there are approximately 50 million epilepsy patients worldwide [1]. Electroencephalogram (EEG) can record neuronal electric activities and is routinely utilized for the diagnosis of epilepsy. Especially prolonged EEG recording is a powerful tool to confirm an epileptic seizure when a patient is manifesting signs and/or symptoms that are difficult judge

from clinical observation, or even when there are no visible symptoms. It is important to recognize seizures early because prolonged seizures can lead to increased brain damage. However, correct interpretation of EEG requires extensive training and EEG specialists may not be readily available in emergency settings. If there is a tool to aid non-expert clinicians to interpret EEG, early recognition and treatment of seizures can eventually improve neurological outcome.

Numerous studies have attempted to detect epileptic seizures quantitatively using EEGs, which are bioelectric signals that can be recorded from the scalp [2]–[8]. For example, Deburachgrave *et al.* detected epileptic seizures based on EEG analysis that focused on the periodicity of high amplitude slow wave activity and sharp changes in amplitude during seizures [2]. Subasi *et al.* also performed a frequency analysis using wavelet transformation to identify epileptic seizures [3]. However, in the above methods, it may be difficult to detect epileptic seizures by evaluating only the magnitude of the amplitude or the changes in frequency, because EEGs recorded during epileptic seizures will differ significantly owing to patient age and the type of epilepsy. In addition, although these methods generally calculate the features and use appropriate discriminators for classification, there are problems such as computational complexity and parameter adjustment. Therefore, in this study, we consider the feature extraction and discriminator problems separately, and focus on the improvement of indicators as the first step to evaluate the discrimination performance by more simple methods such as thresholds.

Meanwhile, the probability distribution shapes of EEGs have been studied as a feature independent of amplitude [9]–[14]. Previous studies typically assumed that EEG signals followed a steady Gaussian distribution [9], [10]; however, after Campbell *et al.* statistically indicated that EEG signals had non-Gaussian properties [11], various studies examined the probability distribution shapes of EEG signals, focusing on their Gaussianity [10], [12]–[14]. Although few previous studies investigated the relationship between epileptic seizures and the Gaussianity of EEG signals, some studies reported that EEG signals during epileptic seizures follow a non-Gaussian distribution [13], [14]. However, these studies are only experimental reports, and currently there is no theoretical framework for quantitatively addressing the non-Gaussianity of EEG signals as a feature.

In this study, we assume that non-Gaussianity in EEG signals is caused by stochastic fluctuations in amplitude, and

*A. Furui, R. Onishi, and *T. Tsuji are with the Graduate School of Engineering, Hiroshima University, Higashi-hiroshima, 739-8527 Japan (e-mail: akirafurui@hiroshima-u.ac.jp; tsuji@bsys.hiroshima-u.ac.jp).

A. Takeuchi is with the Department of Neonatology, Okayama Medical Center, National Hospital Organization.

*T. Akiyama is with the Department of Child Neurology, Okayama University Hospital (e-mail: takiyama@okayama-u.ac.jp).

propose a multivariate scale mixture model that can estimate these fluctuations from recorded EEG signals. In the proposed model, a scalp EEG signal recorded at a certain time using multichannel electrodes follows a multivariate Gaussian distribution, and its variance-covariance matrix is considered as a random variable that follows an inverse Wishart distribution. Thereby, it allows evaluation of the stochastic fluctuation of the variance-covariance matrix corresponding to the EEG amplitude. In addition, by combining this model with a filter bank, we develop an analysis method that can extract time-series stochastic fluctuations in each EEG frequency band as the feature. In the simulation experiment, parameter estimation based on the proposed method is conducted on artificial data, and the error between true values and estimated values observed when conditions are changed is discussed. In the analysis experiment, the proposed method is used to analyze EEG signals recorded during focal seizures. We also verify the goodness-of-fit of the proposed model and investigate the correspondence between (1) estimated stochastic fluctuations of EEG signals and (2) diagnoses by an epileptologist.

The reminder of this paper is organized as follows: Section II outlines the structure of the multivariate scale mixture model, the parameter estimation method, and the EEG analysis method considering frequency characteristics. Section III details the simulation experiments and EEG analysis experiments. Section IV presents the results of these experiments, and Section V provides related discussions. Finally, Section VI discusses the conclusions of this study.

## II. METHODS

### A. Multivariate Scale Mixture Model of Scalp EEG Signals

Fig. 1 shows the stochastic relationship between the recorded EEG signal $\mathbf{x} \in \mathbb{R}^D$ ($D$ is the number of electrodes) and its variance-covariance matrix $\mathbf{\Sigma} \in \mathbb{R}^{D \times D}$ as a graphical model. In the proposed model, $\mathbf{x}$ is handled as a random variable that follows a multivariate Gaussian distribution with a mean of zero and a variance-covariance matrix of $\mathbf{\Sigma}$. The variance-covariance matrix $\mathbf{\Sigma}$ is also a random variable for which the distribution is determined by the degrees of freedom parameter $\nu \in \mathbb{R}^+$ and the scale matrix parameter $\mathbf{\Psi} \in \mathbb{R}^{D \times D}$. Therefore, the variance-covariance matrix $\mathbf{\Sigma}$ is interpreted as a latent variable in the model.

First, the conditional distribution of EEG signal $\mathbf{x}$ given $\mathbf{\Sigma}$ is expressed via the following multivariate Gaussian distribution with a mean vector of zero:

$$p(\mathbf{x}|\mathbf{\Sigma}) = \mathcal{N}(\mathbf{x}|\mathbf{0}, \mathbf{\Sigma})$$
$$= \frac{1}{(2\pi)^{D/2}|\mathbf{\Sigma}|^{1/2}} \exp\left[-\frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{x}\right]. \quad (1)$$

The distribution of the variance-covariance matrix is assumed to obey an inverse Wishart distribution $\mathcal{IW}(\mathbf{\Sigma}; \nu, \mathbf{\Psi})$, which is known as a conjugate prior for the variance-covariance matrix of a multivariate Gaussian distribution [15]:

$$p(\mathbf{\Sigma}) = \mathcal{IW}(\mathbf{\Sigma}; \nu, \mathbf{\Psi})$$
$$= \frac{|\mathbf{\Psi}|^{\frac{\nu}{2}}}{2^{\frac{\nu D}{2}}\Gamma_D\left(\frac{\nu}{2}\right)}|\mathbf{\Sigma}|^{-\frac{\nu+D+1}{2}}\exp\left[-\frac{\mathrm{tr}(\mathbf{\Psi}\mathbf{\Sigma}^{-1})}{2}\right], \quad (2)$$
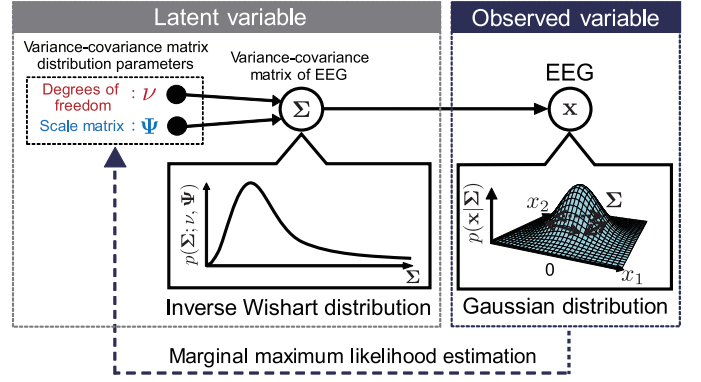


Fig. 1. Graphical representation of the proposed multivariate scale mixture model, which describes the stochastic relationship between EEG signals and its variance-covariance matrix. The white nodes are random variables and the black nodes are parameters to be estimated. In the model, an EEG signal $\mathbf{x} \in \mathbb{R}^D$ is handled as a random variable that follows a multivariate Gaussian distribution with a mean vector of $\mathbf{0}$ and a variance-covariance matrix $\mathbf{\Sigma} \in \mathbb{R}^{D \times D}$ which is also a random variable following the inverse Wishart distribution determined by the degrees of freedom $\nu \in \mathbb{R}^+$ and the scale matrix $\mathbf{\Psi} \in \mathbb{R}^{D \times D}$. The parameter $\nu$ represents stochastic fluctuations of EEG signals. The variance-covariance matrix distribution parameters are estimated via marginal likelihood maximization from recorded EEG signals.

where $\nu$ and $\mathbf{\Psi}$ determine the inverse Wishart distribution, and are referred to as the degrees of freedom and the scale matrix, respectively. Considering the marginal distribution of $\mathbf{x}$, the variance-covariance matrix $\mathbf{\Sigma}$ can be integrated out as follows:

$$p(\mathbf{x}) = \int p(\mathbf{\Sigma})p(\mathbf{x}|\mathbf{\Sigma})\mathrm{d}\mathbf{\Sigma}$$
$$= \int \mathcal{IW}(\mathbf{\Sigma}; \nu, \mathbf{\Psi})\mathcal{N}(\mathbf{x}|\mathbf{0}, \mathbf{\Sigma})\mathrm{d}\mathbf{\Sigma} \quad (3)$$
$$= \int \frac{|\mathbf{\Psi}|^{\frac{\nu}{2}}}{2^{\frac{\nu D}{2}}\Gamma_D\left(\frac{\nu}{2}\right)}|\mathbf{\Sigma}|^{-\frac{\nu+D+1}{2}}\exp\left[-\frac{\mathrm{tr}(\mathbf{\Psi}\mathbf{\Sigma}^{-1})}{2}\right]$$
$$\times \frac{1}{(2\pi)^{D/2}|\mathbf{\Sigma}|^{1/2}}\exp\left[-\frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{x}\right]\mathrm{d}\mathbf{\Sigma}$$
$$= \frac{|\mathbf{\Psi}|^{\frac{\nu}{2}}}{2^{\frac{\nu D}{2}}(2\pi)^{\frac{D}{2}}\Gamma_D\left(\frac{\nu}{2}\right)}\int|\mathbf{\Sigma}|^{-\frac{\nu+D+2}{2}}$$
$$\times \exp\left[-\frac{1}{2}\mathrm{tr}\{(\mathbf{\Psi}+\mathbf{x}\mathbf{x}^{\mathrm{T}})\mathbf{\Sigma}^{-1}\}\right]\mathrm{d}\mathbf{\Sigma}$$
$$= \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu-D+1}{2})}\frac{|\frac{1}{\nu-D+1}\mathbf{\Psi}|^{-\frac{1}{2}}}{[\pi(\nu-D+1)]^{\frac{D}{2}}}(1+\Delta)^{-\frac{\nu+1}{2}}, \quad (4)$$

where $\Delta$ is the square of the Mahalanobis distance, as follows:

$$\Delta = \mathbf{x}^{\mathrm{T}}\mathbf{\Psi}^{-1}\mathbf{x}. \quad (5)$$

From (1) and (3), $p(\mathbf{x})$ is obtained by summing an infinite number of multivariate Gaussian distributions having different variance-covariance matrices, meaning it can be interpreted as the scale mixture of multivariate Gaussian distributions. As stated above, the marginal distribution of multidimensional EEG signals can be modeled based on a multivariate scale mixture distribution.

## B. Parameter Estimation Based on Marginal Maximum Likelihood

Let us consider the estimation of $\nu$ and $\boldsymbol{\Psi}$, given $N$ samples of EEG signals $\mathbf{X} = \{\mathbf{x}_n \in \mathbb{R}^D; n = 1, 2, \cdots, N\}$. The model parameters can be estimated by maximizing the marginal likelihood $p(\mathbf{X}) = \prod_{n=1}^{N} p(\mathbf{x}_n)$. However, obtaining the maximum likelihood solution of a marginal likelihood is generally complex, and thus it is difficult to optimize the solution analytically [15]. Therefore, we conduct this optimization for $\nu$ and $\boldsymbol{\Psi}$ based on the expectation-maximization (EM) algorithm [16].

To simplify (4), we define the new parameters as:

$$\nu = \nu' + D - 1, \tag{6}$$

$$\boldsymbol{\Psi} = \nu' \boldsymbol{\Psi}'. \tag{7}$$

Accordingly, the marginal distribution can be expressed as

$$p(\mathbf{x}_n) = \int \mathcal{IW}(\boldsymbol{\Sigma}_n; \nu' + D - 1, \nu' \boldsymbol{\Psi}') \mathcal{N}(\mathbf{x}_n | \mathbf{0}, \boldsymbol{\Sigma}_n) \mathrm{d}\boldsymbol{\Sigma}_n \tag{8}$$

$$= \frac{\Gamma(\frac{\nu'+D}{2})}{\Gamma(\frac{\nu'}{2})} \frac{|\boldsymbol{\Psi}'|^{-\frac{1}{2}}}{(\pi\nu')^{\frac{D}{2}}} \left(1 + \frac{\Delta'}{\nu'}\right)^{-\frac{\nu'+D}{2}}, \tag{9}$$

where

$$\Delta' = \mathbf{x}_n^{\mathrm{T}} (\boldsymbol{\Psi}')^{-1} \mathbf{x}_n. \tag{10}$$

Equation (9) is equivalent to multivariate Student-$t$ distribution $\mathrm{St}(\mathbf{x}_n | \nu', \boldsymbol{\Psi}')$ [15]. Then, we redefine the latent variable and replace (8) with the following equivalent expression, thereby allowing an efficient calculation (refer to Appendix).

$$p(\mathbf{x}_n) = \int \mathrm{IG}(\tau_n; \nu'/2, \nu'/2) \mathcal{N}(\mathbf{x}_n | \mathbf{0}, \tau_n \boldsymbol{\Psi}') \mathrm{d}\tau_n, \tag{11}$$

where $\tau_n$ is a new latent variable following an inverse Gamma distribution $\mathrm{IG}(\cdot)$. The model parameters $\nu'$ and $\boldsymbol{\Psi}'$ are estimated by maximizing the marginal likelihood as outlined below.

(i) Initialize each parameter by selecting arbitrary starting values.

(ii) *E-step*. Calculate the expectation of the complete-data log-likelihood, denoted as $Q(\nu', \boldsymbol{\Psi}')$.

$$Q(\nu', \boldsymbol{\Psi}')$$
$$= \mathbb{E}\left[\ln \prod_{n=1}^{N} \mathrm{IG}(\tau_n; \nu'/2, \nu'/2) \mathcal{N}(\mathbf{x}_n | \mathbf{0}, \tau_n \boldsymbol{\Psi}')\right]$$
$$= \sum_{n=1}^{N} \left[ -\frac{D}{2} \ln(2\pi) - \frac{D}{2} \mathbb{E}[\ln \tau_n] - \frac{1}{2} \ln |\boldsymbol{\Psi}'| \right.$$
$$- \frac{1}{2} \mathbb{E}[\tau_n^{-1}] \Delta' + \frac{\nu'}{2} \ln \frac{\nu'}{2} - \ln \Gamma\left(\frac{\nu'}{2}\right)$$
$$\left. - \left(\frac{\nu'}{2} + 1\right) \mathbb{E}[\ln \tau_n] - \frac{\nu'}{2} \mathbb{E}[\tau_n^{-1}] \right], \tag{12}$$

where $\mathbb{E}[\tau_n^{-1}]$ and $\mathbb{E}[\ln \tau_n]$ are derived by calculating the posterior distribution $p(\tau_n | \mathbf{x}_n)$ of the latent variable $\tau_n$ as follows:

$$\mathbb{E}[\tau_n^{-1}] = \frac{\nu' + D}{\nu' + \Delta'}, \tag{13}$$

$$\mathbb{E}[\ln \tau_n] = -\ln \mathbb{E}[\tau_n^{-1}] + \ln\left(\frac{\nu'}{2}\right) - \psi\left(\frac{\nu'}{2}\right), \tag{14}$$

where $\psi(\cdot)$ is a digamma function.

(iii) *M-step*. Update the parameters by maximizing $Q(\nu', \boldsymbol{\Psi}')$. By setting the derivative of $Q(\nu', \boldsymbol{\Psi}')$ with $\boldsymbol{\Psi}'$ equal zero, the new scale matrix is obtained as

$$^{\mathrm{new}}\boldsymbol{\Psi}' = \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}[\tau_n^{-1}] \mathbf{x}_n \mathbf{x}_n^{\mathrm{T}}. \tag{15}$$

Because there is no closed-form expression for the degrees of freedom parameter $\nu'$, we estimate $\nu'$ by iteratively maximizing $Q(\nu', \boldsymbol{\Psi}')$ using a bisection method.

$$^{\mathrm{new}}\nu' = \arg\max_{\nu'} Q(\nu', {}^{\mathrm{new}}\boldsymbol{\Psi}'). \tag{16}$$

(iv) Evaluate the log-likelihood $\ln p(\mathbf{X})$ and repeat steps (ii)–(iv) until the calculation converges. Finally, estimated parameters $\nu'$ and $\boldsymbol{\Psi}'$ are transformed to variance-covariance matrix distribution parameters $\nu$ and $\boldsymbol{\Psi}$.

Using these procedures, the parameters of the proposed model can be estimated from recorded EEG signals. Here, the parameter $\nu$ corresponds to the degrees of freedom of the multivariate Student-$t$ distribution parameter (refer to (6)); hence, $\nu$ controls the Gaussianity of a distribution. In the framework of the scale mixture model, the Gaussianity of EEG is considered to change owing to the stochastic fluctuation of the scale parameter (i.e., the variance-covariance matrix). Therefore, the fluctuation of the variance-covariance matrix can be evaluated by estimating $\nu$ from recorded EEG signals.

## C. Proposed EEG Analysis Methods

Fig. 2 shows an overall outline of the proposed analysis method. In the proposed method, observed EEG signals are decomposed into multiple frequency bands by using a filter bank consisting of parallel band-pass filters. Additionally, we estimate non-Gaussianity for signals in each frequency band based on the proposed multivariate scale mixture model.

First, the EEG signal at time $t$ recorded from the $D$ pair of electrodes is defined as $\mathbf{x}_t \in \mathbb{R}^D$. Then, $\mathbf{x}_t$ is divided into $M$ frequency bands $(b_1, \cdots, b_M)$ by applying a filter bank consisting of a third-order Butterworth band-pass filter to the signal; the obtained signal is defined as $x_t^{(b_m)}$ ($m = 1, \cdots, M$). Second, parameter estimation for the variance-covariance matrix distribution based on the proposed model is performed on $\mathbf{x}_t^{(b_m)}$ for each frequency band, and an $\nu$ characterizing the stochastic fluctuation is obtained. Because the characteristics of EEG signals change significantly in a time series, a sliding window of length $W$ (s) is applied to $\mathbf{x}_t^{(b_m)}$ in each band, and $\nu$ is estimated from the sample in the window following the procedure described in Section II-B. The time window is slid by $S$ (s) continuously, resulting in the estimation of $\nu$ in a time series.

Here, $\nu$ is a parameter that controls the Gaussianity of EEG signals. The EEG distribution becomes closer to a Gaussian form as the value of $\nu$ approaches $\infty$; that is, the stochastic fluctuation decreases as $\nu$ increases. In this paper, we calculate
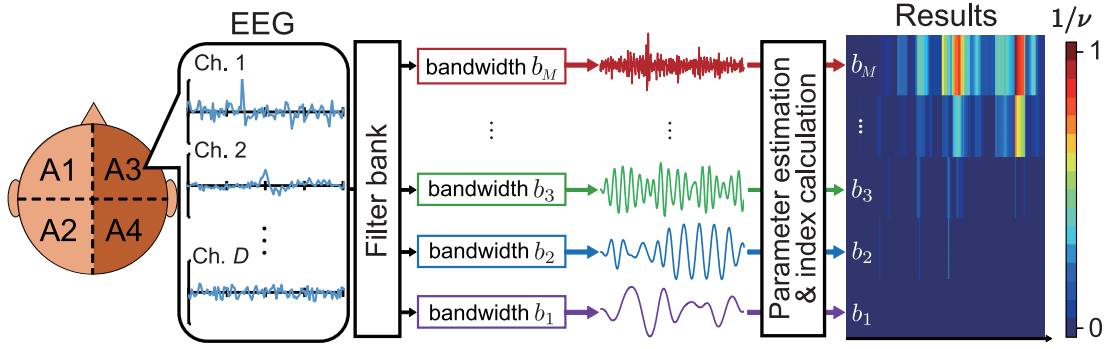
Fig. 2. Overview of the proposed analysis method. The recorded EEG signals are decomposed into multiple frequency bands ($b_1$–$b_M$) by using a filter bank consisting of parallel band-pass filters. The evaluation index $1/\nu$ is then calculated for each frequency band, based on the parameter estimation results of the proposed model. The results are shown as a color map.
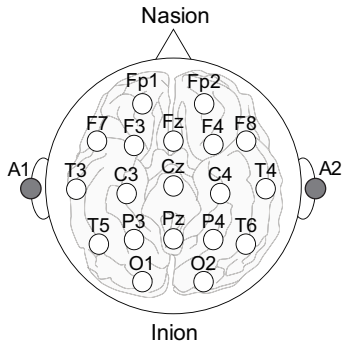


Fig. 3. International 10–20 electrode montage

$1/\nu$, which is the reciprocal of $\nu$, to provide intuitive meaning and to serve as an index that characterizes the non-Gaussianity. The index $1/\nu$ indicates that the larger its value, the larger the stochastic fluctuation of the EEG signal.

From the above, it is clear that the stochastic fluctuations latent in each frequency band of EEG signals can be estimated continuously. It is also possible to obtain the spatial distribution of stochastic fluctuations by dividing the EEG electrode arrangement into multiple regions in advance and performing the above analysis for each region.

## III. EXPERIMENTS

### A. Simulation

To verify the accuracy of parameter estimation based on the proposed model, we performed a simulation experiment to evaluate the error rate between the true and estimated values of the parameters. Using the fact that the marginal distribution of the proposed model is equivalent to the multivariate Student-$t$ distribution, a random number sequence $\{\mathbf{x}_t \in \mathbb{R}^D; t = 1, \cdots, T\}$ following a Student-$t$ distribution $\mathrm{St}(\mathbf{x}_t|\nu_0', \mathbf{\Psi}_0')$ was generated. Here, the $\{\mathbf{x}_t\}$ values were regarded as a time series of an EEG signal recorded at a sampling frequency of $f_s$. The accuracy of the distribution estimation was verified by comparing true values $\nu_0$ and $\mathbf{\Psi}_0$ with estimated values $\nu$ and $\mathbf{\Psi}$, after converting $\nu_0'$ and $\mathbf{\Psi}_0'$ to inverse Wishart distribution parameters based on (6) and (7). As an index of estimation

accuracy for each parameter, the absolute percentage error was defined as $|\nu_0 - \nu|/|\nu_0| \times 100$ and $\|\mathbf{\Psi}_0 - \mathbf{\Psi}\|_F/\|\mathbf{\Psi}_0\|_F \times 100$. $\|\mathbf{\Psi}\|_F$ is the Frobenius norm [17] of $\mathbf{\Psi}$ and is obtained as

$$\|\mathbf{\Psi}\|_F = \sqrt{\sum_{i=1}^{D}\sum_{j=1}^{D} |\psi_{ij}|^2}, \quad (17)$$

where $\psi_{ij}$ is the element of $\mathbf{\Psi}$. In the estimation of each parameter, the first $W$ values of the signals $\{\mathbf{x}_t\}$ were used. The window length $W$ took values of 1, 2, 5, 10, 15, 20, 30, 50, and 100 s, and the number of dimensions $D$ took values of 1, 2, 4, 8, 16, and 19. The average absolute percentage errors were calculated by changing the true values of multivariate $t$-distribution parameters 400 times ($\nu_0' = 0.5, 1.0, 1.5, \cdots, 10.0, \psi_{0ii}' = 1.0, 2.0, 3.0, \cdots, 20.0$). Note that $\mathbf{\Psi}_0'$ was changed only in terms of its diagonal components $\psi_{0ii}'$, and the off-diagonal components were fixed at 0.5. The $T$ and $f_s$ values in artificial data generation were set as 100 s and 500 Hz.

### B. EEG analysis

To evaluate the validity of the proposed model and the effectiveness of the proposed index $1/\nu$ for epileptic seizure detection, analysis experiments were conducted. Twenty epileptic patients with focal epilepsy participated in the experiments. Table I contains patient information, analysis times, and the durations of the epileptic seizures. The EEG signals were recorded with a digital sampling frequency of 500 Hz using an electroencephalograph system (Neurofax EEG-1218, Nihon Kohden, Tokyo, Japan) while the patients were in the supine position. The 19-channel surface electrodes ($D = 19$) were placed on the scalp according to the international 10–20 electrode system, with reference electrodes on both earlobes: A1 and A2 (see Fig. 3). The experiments were approved by the Okayama University Ethics Committee (approval No: 1706-019). The onset and offset of a focal seizure in each EEG recording were marked by a board-certified epileptologist (T.A.).

First, the proposed model was fitted to the recorded EEG signals for all participants, and the index $1/\nu$ was calculated. We set the frequency bands in the proposed method to

TABLE I
PATIENT CONDITIONS

| Patient | Sex | Age (year) | Total data length (s) | Seizure duration (s) |
|---|---|---|---|---|
| A | Male | 2 | 300 | 71 |
| B | Male | 23 | 300 | 54 |
| C | Male | 2 | 300 | 31 |
| D | Male | 4 | 380 | 93 |
| E | Female | 31 | 320 | 39 |
| F | Male | 41 | 300 | 32 |
| G | Female | 3 | 240 | 53 |
| H | Male | 19 | 390 | 36 |
| I | Male | 0.8 | 360 | 98 |
| J | Male | 20 | 390 | 16 |
| K | Male | 36 | 300 | 23 |
| L | Male | 9 | 300 | 43 |
| M | Male | 13 | 300 | 43 |
| N | Male | 15 | 300 | 48 |
| O | Male | 8 | 300 | 17 |
| P | Female | 19 | 300 | 69 |
| Q | Male | 27 | 420 | 62 |
| R | Male | 38 | 300 | 65 |
| S | Male | 17 | 300 | 17 |
| T | Male | 19 | 300 | 65 |



Fig. 4. Examples of artificially generated EEG signals ($D = 4$) with $\nu_0$ set to (a) $\nu_0 = 5.0$, (b) $\nu_0 = 8.0$, and (c) $\nu_0 = 13.0$; in each case, $\psi_{0ii}$ was set to $\psi_{0ii} = 10.0$, $\psi_{0ii} = 20.0$, and $\psi_{0ii} = 40.0$.

$b_m \in \{\delta, \theta, \alpha, \beta, \gamma\}$, and decomposed EEG signals into $\delta$ (1–3 Hz), $\theta$ (4–7 Hz), $\alpha$ (8–12 Hz), $\beta$ (13–24 Hz), and $\gamma$ (25–100 Hz). These frequency bands are generally used to extract features of EEG signals [18]. Here, the fitting of the model and the calculation of $1/\nu$ were performed continuously using a sliding window with a length of $W = 15$ s and a sliding width of $S = 1$ s.

We then performed a goodness-of-fit test to validate the proposed model for EEG signals in each frequency band. As an evaluation index for the goodness-of-fit, we used the Bayesian information criterion (BIC) [19], which balances the fitness and complexity of the model.

$$\text{BIC} = -2\ln L(\hat{\theta}) + k\ln(N_W). \tag{18}$$

Here, $\ln L(\hat{\theta})$ is the log-likelihood of the model, $k$ is the number of estimated parameters in the model, and $N_W$ is the sample size in the sliding window. BIC was calculated for the fitting results for each sliding window. For comparison, BIC was also calculated for a multivariate Gaussian distribution model and a multivariate Cauchy distribution model, which has a heavier tail than the Gaussian distribution.

Next, to verify the effectiveness of the evaluation of EEG non-Gaussianity based on the proposed method, we investigated whether epileptic seizures and non-seizures were classified accurately by using the calculated indices $1/\nu$ of each frequency band ($\delta$–$\gamma$ bands). The calculated $1/\nu$ results for each patient were divided into those obtained from seizure segments and those obtained from non-seizure segments. Here, a non-seizure segment is defined as the period from the start of the recording to the onset of seizures. Non-seizure segments that included extraneous noise owing to unexpected body movements or electrode shifts in the original EEG signals were excluded from the analysis. In addition, beacause the sample size for calculating $1/\nu$ differed significantly between the seizure and non-seizure segments, the calculated $1/\nu$ in the non-seizure segment was randomly sampled based on the
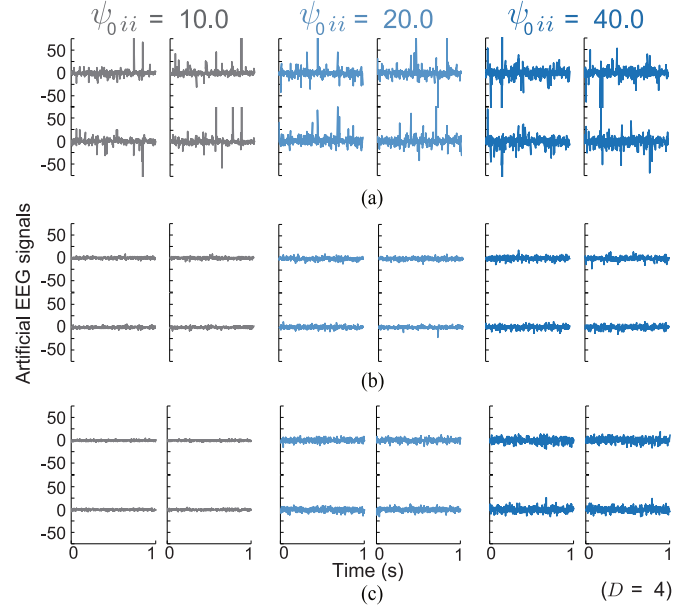
sample size for calculating $1/\nu$ in the seizure segment, and the sample size for calculating $1/\nu$ in both sections was made uniform for each patient.

As an evaluation index of classification performance, we calculated the area under the curve (AUC) based on a receiver operating characteristic (ROC) analysis. The AUC is an evaluation scale calculated from the ROC curve plotting the relationship between the false positive rate and the true positive rate. The closer the AUC value is to 1, the higher the classification performance. For comparison, the AUC was calculated in the same manner using only the EEG amplitude information. The root mean square (RMS) [20] was used for the amplitude information:

$$\text{RMS} = \sqrt{\frac{1}{N_W} \sum_{i=1}^{N_W} (x_i^{\text{Cz}})^2}, \tag{19}$$

where $x_i^{\text{Cz}}$ is an EEG signal at the Cz channel (see Fig. 3) obtained from the top of the head. This RMS was calculated continuously using a sliding window with the same settings as those used in calculating $1/\nu$.

## IV. RESULTS

In the simulation experiments, we generated time-series simulation data and verified the estimation accuracy of the distribution parameter. Fig. 4 shows examples of time-series waveforms of $\mathbf{x}_t$, where the parameters of the inverse Wishart distribution used in artificial data generation were set as $\nu_0 = 5.0$, $\nu_0 = 8.0$, and $\nu_0 = 13.0$, and $\psi_0$ was changed to $\psi_{0ii} = 10.0$, $\psi_{0ii} = 20.0$, and $\psi_{0ii} = 40.0$ for each $\nu_0$. In the examples, the number of dimensions for artificial EEG signals is set to $D = 4$, and the signals for each dimension are shown. The vertical and horizontal axes indicate the signal values and time,
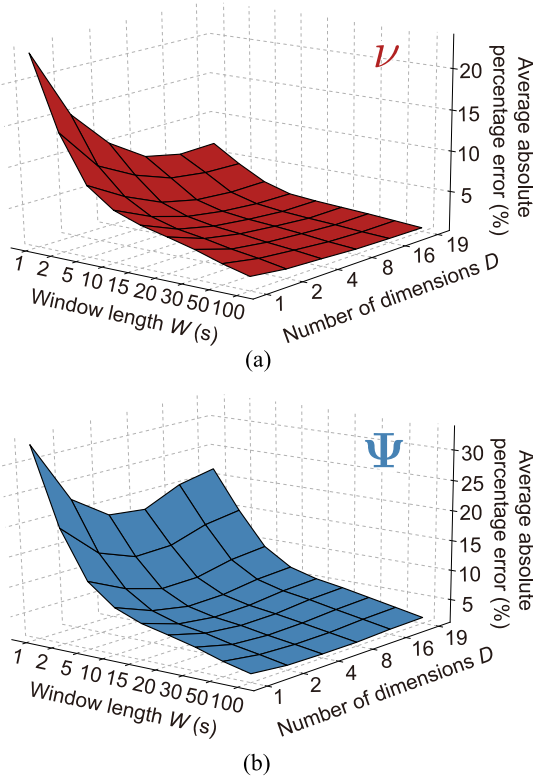
Fig. 5. Average absolute percentage errors for each combination of the number of input dimensions ($D$) and the sliding window length ($W$) in the estimation of variance-covariance distribution parameters. (a) Degrees of freedom $\nu$. (b) Scale matrix $\boldsymbol{\Psi}$.

respectively. Fig. 5 shows the average absolute percentage errors in the estimation of $\nu$ and $\boldsymbol{\Psi}$ resulting from changing the window length $W$ and the number of dimensions $D$.

In the EEG analysis experiments, we performed a goodness-of-fit test for EEG signals and evaluated non-Gaussianity using the proposed index. Table II shows the percentage of times that the BIC of each model became a minimum in the EEG signals of the $\delta$−$\gamma$ bands. The table also shows the results of the McNemar test (significance level: 1%) adjusted by the Holm method using the proposed model as a control group. In all frequency bands, significant differences were observed between the proposed model and the other models ($p < 0.001$). In Fig. 6, color maps show the $1/\nu$ calculations corresponding to the raw EEG waveforms of patient A and patient B. In the color maps, $1/\nu$ was normalized to ensure that the maximum value of all frequency bands would be 1 and the minimum value would be 0. The shaded areas in the waveforms and the areas surrounded by white dotted lines indicate epileptic seizure occurrences diagnosed by an epileptologist.

Fig. 7 shows the $1/\nu$ distribution in each frequency band for all patients, as calculated by kernel density estimation [21] for each seizure and non-seizure segment. The figure also shows the results of the paired $t$-test (significance level: 1%) and the effect size $g$ [22] between the $1/\nu$ distributions from seizure and non-seizure segments. Here, the effect size $g$ is a statistical index indicating the magnitude of the difference between the mean values of the two distributions. In general, $0.2 \leq g < 0.5$

is interpreted as a small effect size, $0.5 \leq g < 0.8$ is a medium effect size, and $0.8 \leq g$ is a large effect size [23].

Fig. 8(a) shows the ROC curves obtained via seizure and non-seizure classification based on the $1/\nu$ in each frequency band and by using RMS as amplitude information. The AUCs obtained from the ROC curves are shown in Fig. 8(b). The AUCs of $1/\nu$ in each frequency band ($\delta$−$\gamma$) were 0.709, 0.647, 0.615, 0.727, and 0.881, respectively, and the AUC of RMS was 0.816. A Delong's test [24] with the Holm adjustment (significance level: 1%) was performed to test the AUCs between the $1/\nu$ in each band and the RMS used as the conventional feature representing the amplitude (i.e., the AUC of RMS was set as a control group). There were significant differences between the AUCs of $1/\nu$ in each band and the AUC of RMS ($p < 0.001$).

## V. DISCUSSION

In the simulation experiment, as the value of $\nu_0$ changes from $\nu_0 = 5.0$ to $\nu_0 = 13.0$, it can be seen that the outlier-like variance of the waveform does not appear frequently, and the waveform stabilized (Fig. 4). This indicates that the distribution approaches a Gaussian distribution as $\nu$ increases, because this parameter determines the Gaussianity of the distribution. It can also be seen that the amplitude of the artificial data increases as the value of $\psi_{0ii}$ changes from $\psi_{0ii} = 10.0$ to $\psi_{0ii} = 40.0$. This is because $\psi_{ii}$, the diagonal component of $\boldsymbol{\Psi}$, is a parameter that characterizes the scale of the variance of EEG signals in each dimension.

The average absolute percentage errors in the estimation of $\nu$ and $\boldsymbol{\Psi}$ are approximately 2% at $W = 100$ s, which indicates that the estimation is accurate (Fig. 5). However, the error rate increases as the window length $W$ decreases; the error rate is approximately 25% in $\nu$ and approximately 35% in $\boldsymbol{\Psi}$ at $W = 1$ s. These results indicate that the estimation accuracy depends on the window length used for parameter estimation. When the number of dimensions was increased, the average absolute percentage error of $\nu$ decreased, and that of $\boldsymbol{\Psi}$ decreased initially and then increased again. Here, $\nu$ is a one-dimensional parameter defined for all input dimensions; therefore, it is considered that increasing the number of dimensions substantially increased the sample size used in the estimation of $\nu$, thereby decreasing the average absolute percentage error in the estimation. The scale matrix $\boldsymbol{\Psi}$ is affected by the estimation result of $\nu'$ because $\boldsymbol{\Psi} = \nu'\boldsymbol{\Psi}'$; therefore, the estimation error of $\boldsymbol{\Psi}$ is considered to decrease synergistically when the number of dimensions increases. Indeed, the estimation error decreased as the number of dimensions ranged from $D = 1$ to $D = 8$, whereas it increased as the number of dimensions ranged from $D = 8$ to $D = 16$. This is probably because the calculation of the average absolute percentage error was performed using the Frobenius norm. Even if the estimation accuracy for each element in the matrix does not change, the overall error increases as the number of elements increases with increases in the number of dimensions.

From the above, it is clear that as the window length increases, the estimation of $\nu$ and $\boldsymbol{\Psi}$ is performed more accurately. As the number of input dimensions increases,
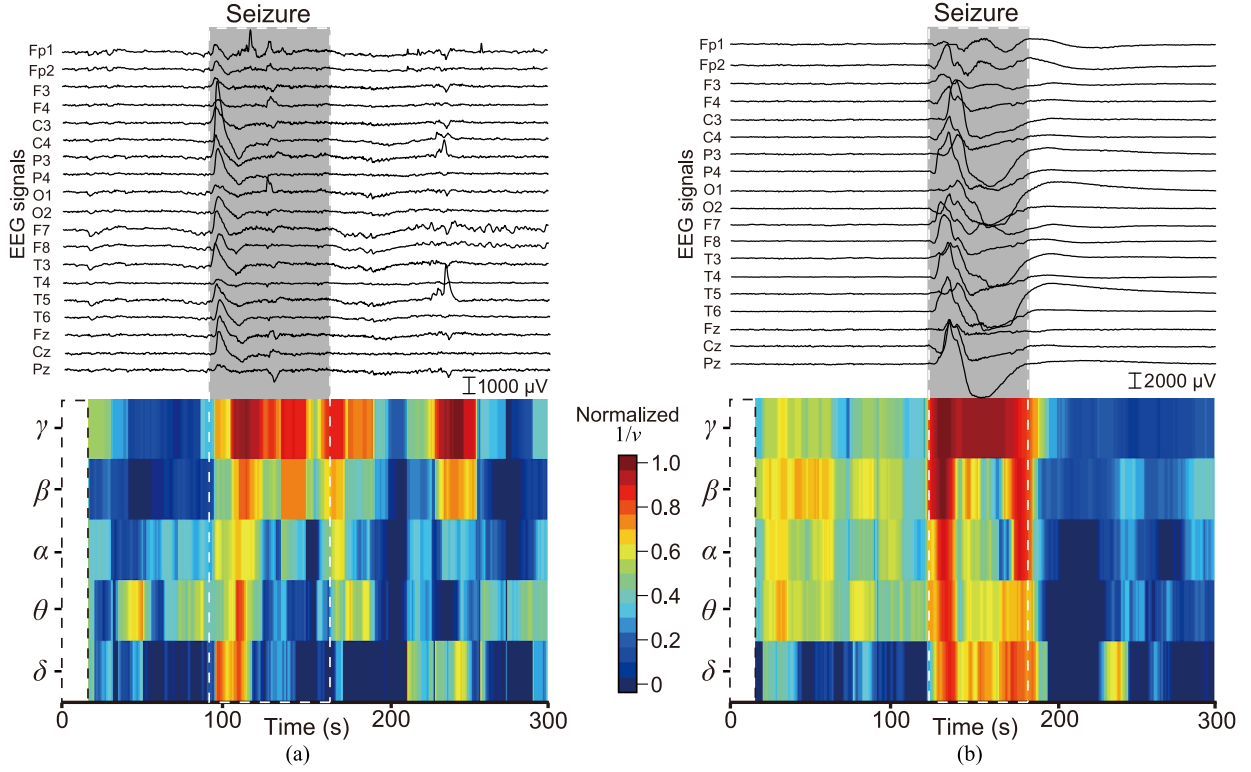
Fig. 6. Raw EEG signals and corresponding analysis results from the proposed method for (a) patient A and (b) patient B. Proposed index $1/\nu$ was normalized based on min-max normalization, rescaling the range of features to scale the range in [0, 1].

TABLE II
PERCENTAGE OF TIMES EACH MODEL WAS SELECTED FOR DIFFERENT
FREQUENCY BANDS BASED ON BIC.

| Frequency band | Model | | |
|---|---|---|---|
| | Proposed | Gaussian | Cauchy |
| $\delta$ (1–3 Hz) | 81.80% | 17.92% *** | 0.28% *** |
| $\theta$ (4–7 Hz) | 95.47% | 4.42% *** | 0.11% *** |
| $\alpha$ (8–12 Hz) | 96.79% | 3.09% *** | 0.12% *** |
| $\beta$ (13–24 Hz) | 99.05% | 0.59% *** | 0.36% *** |
| $\gamma$ (25–100 Hz) | 99.28% | 0.21% *** | 0.51% *** |

***: significant difference with the proposed model as indicated
by a McNemar test with a Holm adjustment ($p < 0.001$)

the estimation of these parameters is also performed more accurately; however, the estimation accuracy of $\Psi$ decreases for higher-dimensional inputs. Additionally, assuming the use of an international 10–20 electrode montage (i.e., the input signal has 19 dimensions), the model parameters can be estimated with an error rate of 10% or less when the window length exceeds 5 s. In particular, setting the window length to 15 s can reduce the error rate to 5%.

The EEG analysis experiment showed that the proposed model is selected the greatest number of times in all frequency bands based on BIC (Table II). However, the ratio of the minimum BIC of the multivariate Gaussian distribution model was 17.92% in the $\delta$ band, which was relatively higher than the others. One possible explanation is that the distribution shape for each window length differs significantly in the low-

frequency bands. In addition, the number of parameters in the proposed model ($k = 382$) is slightly higher than that the multivariate Gaussian distribution model ($k = 381$). Accordingly, when the signal fits both the proposed model and the Gaussian model almost equally, the Gaussian model, a simpler model, shows better goodness-of-fit owing to the nature of BIC. However, in the other frequency bands, the percentage at which the proposed model was selected based on the minimum BIC exceeds 95%. This is because the proposed model can change parameter $\nu$ to adapt to the shape of EEG distributions (i.e., Gaussianity) that change momentarily according to the state of brain activity. In contrast, the Gaussian model and Cauchy model have only a certain Gaussianity, and these two models are included in the proposed model as special cases (Gaussian: $\nu \rightarrow \infty$, Cauchy: $\nu = D$). From the above, we can conclude that the proposed model is more suitable for EEG signals than the other models. In addition, the ratio of the minimum BIC of the proposed model increases as the frequency band becomes higher. This is because the EEG distribution changes depending on frequency characteristics, and those of high-frequency bands (at which signals change rapidly) are better fitted to the proposed model.

In Fig. 6, the $1/\nu$ in the high-frequency band is particularly large in the epileptic seizure segments. This can be confirmed from the results of all patients; therefore, the proposed method can quantitatively evaluate the change in non-Gaussianity during epileptic seizures as stochastic fluctuations. However, at approxiamtely 220 s in Fig. 6(a), a partial increase in $1/\nu$ was observed in intervals other than epileptic seizure segments.
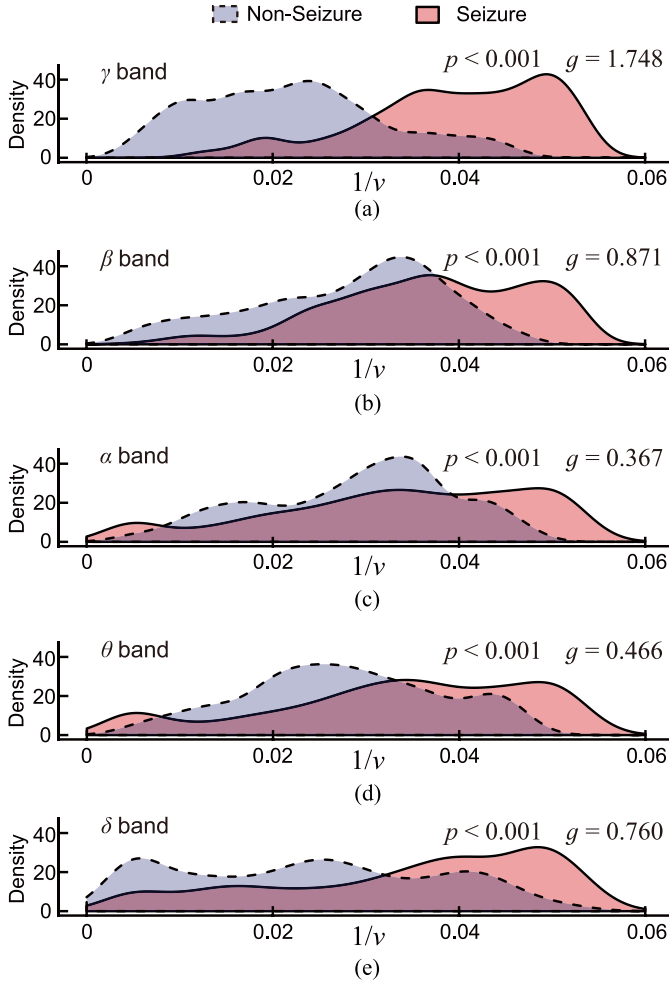
Fig. 7. Experimental $1/\nu$ distributions estimated from EEG signals. The $p$-value from the paired $t$-test and the effect size $g$ are also shown. (a) $\gamma$ band. (b) $\beta$ band. (c) $\alpha$ band. (d) $\theta$ band. (e) $\delta$ band.
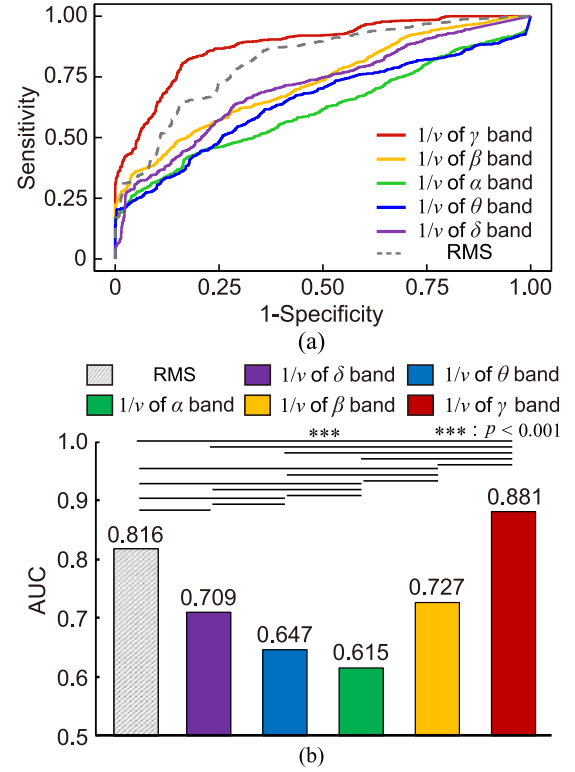


Fig. 8. Results of ROC analysis based on $1/\nu$ for each frequency band and RMS. (a) ROC curves for seizure and non-seizure segments. (b) Calculated AUC for seizure and non-seizure segments. The pairwise comparisons between RMS and the other results based on the Delong's test for two correlated ROC curves with the Holm adjustment are also shown ($p < 0.001$).

This may have occurred because of artifacts in high-frequency bands, e.g., electromyograms caused by body movements. The proposed index, which characterizes stochastic fluctuations, may increase to some extent owing to non-stationary noise components superimposed onto the EEG.

In Fig. 7, the $1/\nu$ distributions of all patients in each frequency band show that the $1/\nu$ of non-epileptic seizure segments is distributed in a region smaller than 0.05 regardless of the frequency band. In contrast, the $1/\nu$ of epileptic seizure segments is distributed more widely than that of non-epileptic seizure segments, and this tendency is most noticeable in the high-frequency $\gamma$ band. This can also be confirmed from the effect size $g$, which is the standardized mean difference between two groups. It is considered that the $1/\nu$ in the high-frequency band best reflects the characteristics of epileptic seizures because a large effect size was obtained in the $\gamma$ band. This indicates that EEG signals in the high-frequency band exhibited a strong non-Gaussianity owing to epileptic seizures. Previous studies reported that the activity in the $\gamma$ band, which is the high-frequency band of EEG signals, becomes intense during epileptic seizures [25]–[27]. The de-

tailed mechanism of such $\gamma$ activity is still unknown; however, activities of inhibitory interneurons and electrical coupling through gap junctions were previously suggested [28]. These activities cause intermittent amplitude changes specific to epileptic seizures; consequently, the non-Gaussianity in the high-frequency band may be emphasized and $1/\nu$ increases.

In Fig. 8(b), a medium to high AUC of 0.881 was obtained in the high-frequency $\gamma$ band, showing that $1/\nu$ in the $\gamma$ band had the highest classification performance. In addition, significant differences were noted with the amplitude feature (i.e., RMS); therefore, the $1/\nu$ in the $\gamma$ band provided better classification ability than the conventional feature. It is considered that the scale matrix $\Psi$, which includes amplitude information, and the proposed index $1/\nu$ reflect the different aspects of EEG activity. Thus, if we could combine the features defined in the proposed model, higher classification performance would be obtained.

## VI. CONCLUSION

In this paper, we modeled EEG signals using a multivariate scale mixture model. In the model, a scalp EEG signal recorded at a certain time using multichannel electrodes follows a multivariate Gaussian distribution, and its variance-covariance matrix is handled as a random variable that follows an inverse Wishart distribution. The fluctuation of the variance-covariance matrix can be evaluated based on this stochastic relationship. We also proposed an EEG analysis method by

combining the scale mixture model and a filter bank, and introduced an index $1/\nu$ characterizing the non-Gaussianiy. In this method, EEG signals are decomposed into several frequency bands, and the index for each frequency band was calculated in time series based on the sliding window.

In the simulation experiment, we evaluated the estimation accuracy of each parameter, and showed that the estimation accuracy changes depending on the sample size and the number of dimensions. In experiments using EEG signals that included epileptic seizures, it was demonstrated that the proposed model is the most suitable model for EEGs. In addition, higher accuracy (AUC = 0.881) in classifying seizure and non-seizure segments was obtained by focusing on the proposed index $1/\nu$ in the $\gamma$ band.

However, the proposed index $1/\nu$ may be affected by artifacts such as electromyograms caused by body movement or stiffness owing to seizures. Therefore, to apply the proposed method more practically, it will be necessary to introduce an algorithm that detects and removes these artifacts. In addition, we plan to evaluate another parameter of the inverse Wishart distribution, $\boldsymbol{\Psi}$, and examine its application to EEGs other than EEGs for the diagnosis of epilepsy.

TABLE III
AUC OF ALL FEATURES FOR DIFFERENT FREQUENCY BANDS.

| Frequency band | Features | | | |
|---|---|---|---|---|
| | $1/\nu$ | RMS | \|ToC\| | ApEn |
| $\delta$ (1–3 Hz) | 0.709 | 0.721 | 0.757 | 0.560 |
| $\theta$ (4–7 Hz) | 0.647 | 0.746 | 0.717 | 0.557 |
| $\alpha$ (8–12 Hz) | 0.615 | 0.817 | 0.773 | 0.555 |
| $\beta$ (13–24 Hz) | 0.727 | 0.822 | 0.786 | 0.548 |
| $\gamma$ (25–100 Hz) | 0.881 | 0.878 | 0.853 | 0.552 |

## APPENDIX
## EQUIVALENT EXPRESSION FOR MULTIVARIATE SCALE MIXTURE MODEL

This appendix shows that the equivalent of the (9) and (11). Equation (11) can be calculated as

$$
\begin{aligned}
p(\mathbf{x}_n) \\
&= \int \mathrm{IG}(\tau_n; \nu'/2, \nu'/2) \mathcal{N}(\mathbf{x}_n | \mathbf{0}, \tau_n \boldsymbol{\Psi}') \mathrm{d}\tau_n \\
&= \int \frac{\left(\frac{\nu'}{2}\right)^{\frac{\nu'}{2}}}{\Gamma\left(\frac{\nu'}{2}\right)} (\tau_n)^{-\frac{\nu'}{2}-1} \exp\left[-\frac{1}{\tau_n}\left(\frac{\nu'}{2}\right)\right] \\
&\quad \times \frac{1}{(2\pi)^{\frac{D}{2}} (\tau_n)^{\frac{D}{2}} |\boldsymbol{\Psi}'|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}\mathbf{x}_n^{\mathrm{T}}(\tau_n \boldsymbol{\Psi}')^{-1}\mathbf{x}_n\right] \mathrm{d}\tau_n \\
&= \frac{1}{(2\pi)^{\frac{D}{2}}} \frac{\left(\frac{\nu'}{2}\right)^{\frac{\nu'}{2}}}{\Gamma\left(\frac{\nu'}{2}\right)} \frac{1}{|\boldsymbol{\Psi}'|^{\frac{1}{2}}} \\
&\quad \times \int (\tau_n)^{-\frac{\nu'+D}{2}-1} \exp\left[-\frac{1}{\tau_n}\left(\frac{\nu'+\Delta'}{2}\right)\right] \mathrm{d}\tau_n \\
&= \frac{1}{(2\pi)^{\frac{D}{2}}} \frac{\left(\frac{\nu'}{2}\right)^{\frac{\nu'}{2}}}{\Gamma\left(\frac{\nu'}{2}\right)} \frac{1}{|\boldsymbol{\Psi}'|^{\frac{1}{2}}} \frac{\Gamma\left(\frac{\nu'+D}{2}\right)}{\left(\frac{\nu'+\Delta'}{2}\right)^{\frac{\nu'+D}{2}}} \\
&\quad \times \int \frac{\left(\frac{\nu'+\Delta'}{2}\right)^{\frac{\nu'+D}{2}}}{\Gamma\left(\frac{\nu'+D}{2}\right)} (\tau_n)^{-\frac{\nu'+D}{2}-1} \\
&\quad \times \exp\left[-\frac{1}{\tau_n}\left(\frac{\nu'+\Delta'}{2}\right)\right] \mathrm{d}\tau_n \\
&= \frac{1}{(2\pi)^{\frac{D}{2}}} \frac{\left(\frac{\nu'}{2}\right)^{\frac{\nu'}{2}}}{\Gamma\left(\frac{\nu'}{2}\right)} \frac{1}{|\boldsymbol{\Psi}'|^{\frac{1}{2}}} \frac{\Gamma\left(\frac{\nu'+D}{2}\right)}{\left(\frac{\nu'+\Delta'}{2}\right)^{\frac{\nu'+D}{2}}} \\
&\quad \times \int \mathrm{IG}(\tau_n; \frac{\nu'+D}{2}, \frac{\nu'+\Delta'}{2}) \mathrm{d}\tau_n.
\end{aligned}
\tag{20}
$$

Here, the integral of the probability density function over the entire space is equal to 1. Hence, (20) can be expressed in the same form as (9).

$$
p(\mathbf{x}_n) = \frac{\Gamma(\frac{\nu'+D}{2})}{\Gamma(\frac{\nu'}{2})} \frac{|\boldsymbol{\Psi}'|^{-\frac{1}{2}}}{(\pi\nu')^{\frac{D}{2}}} \left(1 + \frac{\Delta'}{\nu'}\right)^{-\frac{\nu'+D}{2}}.
\tag{21}
$$

tmp

tmp: Feature extraction computation time was also evaluated. The computation time in Debug C++ code with Microsoft Visual Studio with Intel® CoreTM i7–6900K CPU @ 3.2GHz
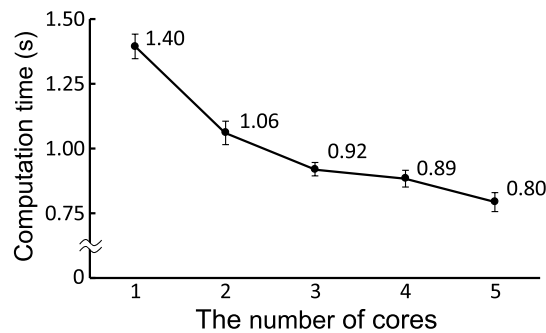
Fig. 9. The average computation time and its standard error of the proposed index $1/\nu$. The sample length took 7500 ($W$=15), and the nmber of dimensions took $D$=19 by changing the number of cores 1–5.

with 64GB RAM was calculated. The computation time was calculated 10 times for the sample length of 7500 for 5 frequency bands in parallel computing by changing the number of cores and the average time and its standard error are shown in Fig. 9.

## REFERENCES

[1] World Health Organization (2019). Epilepsy fact sheet. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/epilepsy

[2] W. Deburchgraeve *et al.*, "Automated neonatal seizure detection mimicking a human observer reading EEG," *Clin. Neurophysiol.*, vol. 119, no. 11, pp. 2447–2454, Jul. 2008.

[3] A. Subasi and E. Ercelebi, "Classification of EEG signals using neural network and logistic regression," *Comput. Meth. Prog. Bio.*, vol. 78, no. 2, pp. 87–99, May 2005.

[4] P. Kellaway *et al.*, "Precise characterization and quantification of infantile spasms," *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, vol. 6, no. 3, pp. 214–218, Sep. 1979.

[5] F. Panzica *et al.*, "Spectral properties of EEG fast activity ictal discharges associated with infantile spasms," *Clin. Neurophysiol.*, vol. 110, no. 4, pp. 593–603, Sep. 1999.

[6] L. Fusco and F. Vigevano, "Ictal clinical electroencephalographic findings of spasms in West syndrome," *Epilepsia*, vol. 34, no. 4, pp. 671–678, Jul. 1993.

[7] N. Acir *et al.*, "Automatic detection of epileptiform events in EEG by a three-stage procedure based on artificial neural networks," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 1, pp. 30–40, Feb. 2005.

[8] B. R. Greene *et al.*, "A comparison of quantitative EEG features for neonatal seizure detection," *Clin. Neurophysiol.*, vol. 119, no. 6, pp. 1248–1261, Feb. 2008.

[9] M. G. Saunders, "Amplitude probability density studies on alpha and alpha-like patterns," *Clin. Neurophysiol.*, vol. 15, no. 5, pp. 761–767, Oct. 1963.

[10] F. F. Gonen and G. V. Tcheslavski, "Techniques to assess stationarity and Gaussianity of EEG: An overview," *Int. J. Bioautom.*, vol. 16, no. 2, pp. 135–142, Jul. 2012.

[11] J. Campbell *et al.*, "On the sufficiency of autocorrelation functions as EEG descriptors," *IEEE Trans. Biomed. Eng.*, vol. BME-14, no. 1, pp. 49–52, Jan. 1967.

[12] M. S. Weiss, "Non-Gaussian properties of the EEG during sleep," *Electroenceph. Clin. Neurophysiol.*, vol. 34, no. 2, pp. 200–202, Aug. 1973.

[13] M. Nurujjaman *et al.*, "Comparative study of nonlinear properties of EEG signals of normal persons and epileptic patients," *Nonlin. Biomed. Phys.*, vol. 3, pp. 1–5, Jan. 2009.

[14] P. J. Charles *et al.*, "Non-Gaussian modeling of EEG data," in *Proc. of the First Joint BMES/EMBS Conf.*, no. 2, Oct. 1999, p. 1023.

[15] C. M. Bishop, *Pattern recognition and machine learning*. New York, USA: Springer-Verlag, 2006.

[16] J. A. Bilmes, "A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models," *Int. Comput. Sci. Inst.*, vol. 4, no. 510, p. 126, Apr. 1998.

[17] G. H. Golub and C. F. Van Loan, *Matrix computations*, 4th ed. Baltimore, MD: The John Hopkins University Press, 2013.

[18] J. Hughes, *EEG in clinical practice*. MA, Boston: Butterworth-Heinemann, 1994.

[19] G. Schwarz, "Estimating the dimension of a model," *Ann. Statist.*, vol. 6, no. 2, pp. 461–464, Mar. 1978.

[20] M. Hamedi *et al.*, "Neural network-based three-class motor imagery classification using time-domain features for BCI applications," *IEEE Region 10 Symp.*, pp. 204–207, Apr. 2014.

[21] E. Parzen, "On estimation of a probability density function and mode," *Ann. Math. Statist.*, vol. 33, no. 3, pp. 1065–1076, Sep. 1962.

[22] L. V. Hedges, "Distribution theory for Glass's estimator of effect size and related estimators," *J. Educational Statist.*, vol. 6, no. 2, pp. 107–128, Jun. 1981.

[23] J. Cohen, *Statistical power analysis for the behavioral sciences*. Evanston, IL, USA: Routledge, 2013.

[24] E. R. DeLong *et al.*, "Comparing the areas under two or more correlated receiver operating characteristic curves : A nonparametric approach," *Biometrics*, vol. 44, no. 3, pp. 837–845, Sep. 1988.

[25] K. Kobayashi *et al.*, "Very fast rhythmic activity on scalp EEG associated with epileptic spasms," *Epilepsia*, vol. 45, no. 5, pp. 488–496, May 2004.

[26] K. Kobayashi *et al.*, "Spectral analysis of EEG gamma rhythms associated with tonic seizures in Lennox-Gastaut syndrome," *Epilepsy Research*, vol. 86, no. 1, pp. 15–22, Sep. 2009.

[27] K. Benedek *et al.*, "Neocortical gamma oscillations in idiopathic generalized epilepsy," *Epilepsia*, vol. 57, no. 5, pp. 796–804, Mar. 2016.

[28] S. C. Park *et al.*, "Ictal high-gamma oscillation (60–99 Hz) in intracranial electroencephalography and postoperative seizure outcome in neocortical epilepsy," *Clin. Neurophysiol.*, vol. 123, no. 6, pp. 1100–1110, Jun. 2012.