International Journal of Neural Systems, Vol. 27, No. 8 (2017) 1750046 (16 pages)
© World Scientific Publishing Company
DOI: 10.1142/S0129065717500460

World Scientific
www.worldscientific.com

EEG Classification with a Sequential Decision-Making Method in Motor Imagery BCI

Rong Liu*
Biomedical Engineering Department
Dalian University of Technology
Dalian, Liaoning 116024, P. R. China
rliu@dlut.edu.cn

Yongxuan Wang
Affiliated Zhongshan Hospital of Dalian University
Dalian University of Technology
Dalian, Liaoning 116001, P. R. China
wangyongxuan@dlu.edu.cn

Geoffrey I. Newman† and Nitish V. Thakor‡
Department of Biomedical Engineering
School of Medicine, Johns Hopkins University
Baltimore, MD 21205, USA
†*geoffrey.newman@gmail.com*
‡*nitish@jhu.edu*

Sarah Ying
Departments of Radiology, Neurology, and Ophthalmology
School of Medicine, Johns Hopkins University
Baltimore, MD 21205, USA
sying@dizzy.med.jhu.edu

Accepted 28 August 2017
Published Online 17 October 2017

Int. J. Neur. Syst. 2017.27. Downloaded from www.worldscientific.com
by ZHEJIANG UNIVERSITY on 11/29/17. For personal use only.

To develop subject-specific classifier to recognize mental states fast and reliably is an important issue in brain–computer interfaces (BCI), particularly in practical real-time applications such as wheelchair or neuroprosthetic control. In this paper, a sequential decision-making strategy is explored in conjunction with an optimal wavelet analysis for EEG classification. The subject-specific wavelet parameters based on a grid-search method were first developed to determine evidence accumulative curve for the sequential classifier. Then we proposed a new method to set the two constrained thresholds in the sequential probability ratio test (SPRT) based on the cumulative curve and a desired expected stopping time. As a result, it balanced the decision time of each class, and we term it balanced threshold SPRT (BTSPRT). The properties of the method were illustrated on 14 subjects' recordings from offline and online tests. Results showed the average maximum accuracy of the proposed method to be 83.4% and the average decision time of 2.77 s, when compared with 79.2% accuracy and a decision time of 3.01 s for the sequential Bayesian (SB) method. The BTSPRT method not only improves the classification accuracy and decision speed comparing with the other nonsequential or SB methods, but also provides an explicit relationship between stopping time, thresholds and error, which is important for balancing the speed-accuracy tradeoff. These results suggest that BTSPRT would be useful in explicitly adjusting the tradeoff between rapid decision-making and error-free device control.

Keywords: Brain–computer interface (BCI); motor imagery; classification; decision-making.

*Corresponding author.

Int. J. Neur. Syst. 2017.27. Downloaded from www.worldscientific.com
by ZHEJIANG UNIVERSITY on 11/29/17. For personal use only.

1. Introduction

Brain–computer interface (BCI) has gained popularity for its different applications, from clinical applications such as a means of communication for locked-in or paralyzed patients, to entertainment, including a new way for gaming.[1–4] To realize intuitive and natural control for many BCI tasks, single-trial classification of EEG data with fewer channels during motor imagery received much attention.[5,39] However, the analysis of single-trial EEG generally suffers from low signal-to-noise ratio (SNR) and must be made robust to adapt nonstationary changes.

The term "classification" in BCI implies decision-making.[6,7] One popular mathematical framework to interpret decision-making is signal detection theory (SDT). It specifies how a single-trial observation converts into a categorical response. However, the evaluation criterion of the SDT is mainly based on the associated classification error.[8] Most SDT methods use a fixed data size which is far from optimal as the elapsed time to reach a decision is also a very important property of decisions. Another simple but powerful idea to interpret decision-making is sequential analysis (SA). In SA, evidence is accumulated across time until reaching to some threshold level, whence the decision is made. Specifically, great progress has been made in characterizing the mechanism of evidence accumulation and optimal decision-making.[9–11]

To control brain-actuated devices, such as robotics and neuro-prostheses, it is important to not only achieve reliable classification (i.e. a minimal error rate) but also to speed up the decision-making process.[12–15] This may pose an inherent tradeoff between classification accuracy and speed of decision-making. This idea of sequential evidence accumulation motivated us to study whether or not a variable-length sequential sampling model would allow us formally and rigorously to manage this tradeoff between accuracy/error rate and decision speed in practical BCI scenarios.

Three kinds of approaches have been developed to use the temporal information in the BCI applications.[6,16,17] One approach is to concatenate features extracted from several time segments into a single vector. This method relies on finding an optimal time segment window for classification. It is commonly applied in BCI applications.[18] A second approach extracts and classifies the features on several time segments and then combines the results from different segments. Mostly, it is used in a rather simple setting by just adding up the classifier outputs over time and setting a threshold for the BCI application. This "moving cursor" method in fact suffers a lot from the nonmonotonicity in practice. In addition, Lemm *et al.*[19] suggested a sequential Bayesian (SB) procedure, which combines the posterior probabilities of a sequence of classifiers to recognize motor imagery tasks. More recently, Gangadhar *et al.*[20] developed a time aggregation of classification method for the recognition of anticipation-related EEG in a contingent negative variation paradigm. These approaches were implemented in a Bayesian framework and have shown improved BCI performance by combining EEG evidence over time. However, they failed to give an explicit relationship between stopping time, thresholds and error probabilities; such a relationship is important to achieve the desired speed-accuracy tradeoff and implement a real-time BCI.

The last method discussed here is a dynamic classification method, which incorporates a temporal sequence of feature vectors into evidence accumulation across time until it makes a decision. For a sequential sampling model, there are two elements, the terminal rule and the decision rule. A systematic theory of computing theoretical optimal stopping time emerged with the work by Wald on the sequential probability ratio test (SPRT), which could achieve a required error rate with fewer numbers of observations on average.[21] The sequential decision-making methods have been tentatively used in BCI research with promising results.[17,22,23] However, the two uncorrelated thresholds in traditional SPRT are often approximated by simple expression of user-chosen error probabilities which are not known *a priori*. Therefore, in this study, we first developed a user-specific wavelet method to determine evidence accumulation curve based on a grid-search method with the training data. Then the link between a decision rule and a termination time was exploited. We built a new flexible framework based on the cumulative curve in conjunction with optimal stopping time to obtain two constrained thresholds. As a result, it balanced the decision time of each class and then achieved a balance between accuracy and decision speed balance.

Int. J. Neur. Syst. 2017.27. Downloaded from www.worldscientific.com
by ZHEJIANG UNIVERSITY on 11/29/17. For personal use only.

This paper provides a novel way to determine the thresholds without predefined error probabilities which are important for practical BCI applications. This work could lead to improved real-world usability for BCI systems.

2. Methodology

2.1. *Data description*

The EEG data used in this work were obtained from 14 subjects. Four subjects were taken from BCI competitions II&III. The remaining 10 were recorded in the Biomedical Engineering Laboratory at Dalian University of Technology (DLUT). The task was performed based on left and right hand motor imagination.

2.1.1. *Dataset* III *from BCI competition* II

This dataset contains EEG data from one subject (S2003).[24] The data were recorded from three electrodes at 128 Hz: C3, Cz and C4. The data consist of 280 trials (140 labeled and unlabeled trials) with equal number of left- and right-hand class. The unlabeled trials were used for testing, whereas the labeled for the training. In each trial, a left or right arrow was presented as a cue after 3 s preparation period. At the same time, the subject did motor imagination task. Each trial lasted 9 s.

2.1.2. *Dataset* III*b from BCI competition* III

The second dataset contains EEG data recorded over the channels C3 and C4 from three subjects (O3, S4, X11) with some corrections.[25] The data were sampled at 125 Hz. Training and testing sets were available for each subject. For subject O3, the remaining data contain 160 training trials and 154 testing trials. In addition, for subjects S4 and X11, there are 494 and 537 training trials, 492 and 539 testing trials, respectively. Each trial has a duration of 7 s which consists of 3 s for preparation period, 1 s for a visual cue presentation and followed by another 3 s for the imagination task.

2.1.3. *Dataset from DLUT*

We supplemented the two BCI competition datasets with our own data recorded with a QuickCap 40-electrode scalp cap in a similar experimental paradigm as previous two datasets. The data were sampled at 125 Hz. Each trial lasted 6 s. In each trial, the subject was instructed to perform motor imagination based on the direction of the arrow after 2 s preparation period. Ten healthy subjects (L1–L10) without prior BCI experience participated in the online experiments. Each subject completed 100 trials. Fifty trials were used for online training of the classifiers, and the remaining 50 were used to determine classification accuracy.

2.2. *Subject-specific feature extraction*

When performing motor imagination, the amplitude changes in the μ (8–13 Hz) and β (18–30 Hz) rhythms are found to reveal event-related desynchronization and synchronization (ERD/ERS) over sensorimotor cortex.[26] Due to the main problem associated with such features and nonstationary effects of brain signals, a wavelet transform is often applied to the EEG signals analysis.[19,27,28] In this paper, we also used the wavelet method to extract feature. The wavelet $\psi(t)$ can be expressed as

$$\psi(t) = c \cdot e^{-\frac{1}{2}\cdot\left(\frac{2\pi ft}{K}\right)^2} \cdot e^{j2\pi ft}, \tag{1}$$

where f is the central frequency of the wavelet, K is the modulation parameter with the restriction $K \geq 5$ (the wavelet admissible condition), j is the complex number and the coefficient c is to ensure $\|\psi(t)\| = 1$.

Instead of using fixed empirical parameters, we found an optimal (f, K) pair by maximizing a constructed t-statistics of each subject to get the wavelet basis $\psi_{f,K}(t)$. That is to say, for a given (f, K) pair, we calculated the band power for both left-hand and right-hand motor imagery EEG across all the time. This produced two sets of data for the t-statistics computation. Then a grid search method was used to obtain the optimal combination of parameters f, K corresponding to the maximum t-statistics.

Let $x_c(t),\ c \in \{C3, C4\}$ denote the EEG recorded over C3 and C4, the projection of the signals to the wavelet is

$$z_{cf} = \langle x_c(t), \psi_f(t) \rangle, \quad c \in \{C3, C4\}, \quad f \in \{\mu, \beta\}. \tag{2}$$

The resulting band power vector is

$$\mathbf{z} = [|z_{C3\mu}|^2, |z_{C3\beta}|^2, |z_{C4\mu}|^2, |z_{C4\beta}|^2]^{\mathrm{T}}. \tag{3}$$

Int. J. Neur. Syst. 2017.27. Downloaded from www.worldscientific.com by ZHEJIANG UNIVERSITY on 11/29/17. For personal use only.

The obtained feature vectors of the training dataset are further used to estimate the parameters of the two class probabilistic models, such as the mean vectors $\mathbf{m}_k$ and the covariance matrices $\mathbf{S}_k$ with $k \in \{L, R\}$ which represents the left- or right-hand motor imagery. With the Gaussian distribution assumption, the probability density function (PDF) is

$$\begin{aligned} g_k(\mathbf{z}) &= g(\mathbf{z} \,|\, \mathbf{m}_k, \mathbf{S}_k) \\ &= (2\pi)^{-2} |\mathbf{S}_k^{-1}|^{\frac{1}{2}} \\ &\quad \cdot \exp\left(-\frac{1}{2}(\mathbf{z} - \mathbf{m}_k)^{\mathrm{T}} \mathbf{S}_k^{-1} (\mathbf{z} - \mathbf{m}_k)\right). \end{aligned} \tag{4}$$

2.3. *Balanced threshold SPRT (BTSPRT) classification method*

The SPRT is a statistical decision model that assumes decisions are formed by continuously sampling information until a response criterion is satisfied.[21,23] The number of observations needed for a decision is not determined in advance of the experiment, but fed to the SPRT algorithm sequentially during the test. Therefore, each trial was separated into segments with the same length as window size of the Morlet wavelet feature length.

Taking into account the nonstationarity of the EEG sampling information, we assume that the probability distribution of the ith segment feature $\mathbf{z}_i$ for class k is $g_k(\mathbf{z}_i), k \in \{L, R\}$, where L and R represent left- or right-hand movement imagination. According to SPRT,[21] the accumulated likelihood ratio $l_q(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_q)$ of sequence $\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_q$ is computed as:

$$l_q(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_q) = \prod_{i=1}^{q} \frac{g_{Li}(\mathbf{z}_i)}{g_{Ri}(\mathbf{z}_i)}, \tag{5}$$

where q is the number of accumulated segments. Assuming that the segments are independent for computation convenience,[21,22] we have

$$l_q(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_q) = \frac{g_L(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_q)}{g_R(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_q)}, \tag{6}$$

where $g_k(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_q), k \in \{L, R\}$ is the join probability distribution of q dimensional vector $(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_q)$. This assumption is violated by our data in practice.

The decision rule with two thresholds η_L and η_R is

$$\begin{cases} \text{if } l_q(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_q) \geq \eta_L, & \hat{k} = L \\ \text{if } l_q(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_q) \leq \eta_R, & \hat{k} = R \\ \text{if } \eta_R < l_q(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_q) < \eta_L, & \text{compute } l_{q+1} \end{cases}.$$

With two thresholds, we have the option to increase η_L or decrease η_R which will increase the probability to make a correct decision (by waiting to accumulate more data or evidence), but decreases the probability of making a wrong decision (by delaying the decision). The error probabilities are defined as

$$\begin{aligned} p_{L|R} &= P(\hat{k} = L \,|\, k = R) \\ p_{R|L} &= P(\hat{k} = R \,|\, k = L) \end{aligned}. \tag{7}$$

If $l_q(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_q) \geq \eta_L$ is satisfied, we define the corresponding space of vector $(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_q)$ to be Z_L. With Eq. (6), we have

$$g_L(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_q) \geq \eta_L \cdot g_R(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_q). \tag{8}$$

Likewise, Eq. (8) is then integrated in the Z_L yields

$$\int_{Z_L} g_L(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_q) \geq \eta_L \cdot \int_{Z_L} g_R(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_q). \tag{9}$$

That is

$$1 - p_{R|L} \geq \eta_L \cdot p_{L|R}. \tag{10}$$

Analogous reasoning for $l_q(\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_q) \leq \eta_R$ yields

$$p_{R|L} \leq \eta_R \cdot (1 - p_{L|R}). \tag{11}$$

Thus, the two detection thresholds η_L and η_R are related to the error probabilities by

$$\begin{aligned} p_{L|R} &\leq \frac{1 - p_{R|L}}{\eta_L} \leq \frac{1}{\eta_L} \\ p_{R|L} &\leq (1 - p_{L|R}) \cdot \eta_R \leq \eta_R \end{aligned}. \tag{12}$$

The two kinds of error probabilities can be lowered by either increasing η_L or decreasing η_R. However, due to the limited number of segments, the indecision rate will increase as the error probability $p_{L|R}$ or $p_{R|L}$ is decreased. So the appropriate selection of η_L and η_R should be essential.

Under the assumption that the features follow a Gaussian distribution, we can take logarithm on both

Int. J. Neur. Syst. 2017.27. Downloaded from www.worldscientific.com by ZHEJIANG UNIVERSITY on 11/29/17. For personal use only.

sides of Eq. (6) to obtain the log probability ratio (log PR), which leads an SPRT as

$$L_q = \sum_{i=1}^{q} \frac{1}{2}\left(\log\frac{|\mathbf{S}_{Ri}|}{|\mathbf{S}_{Li}|} + D_M^2(\mathbf{z}_i, \mathbf{m}_{Ri}) - D_M^2(\mathbf{z}_i, \mathbf{m}_{Li})\right) \triangleq \sum_{i=1}^{q} J_i, \tag{13}$$

where $D_M(\mathbf{z}_i, \mathbf{m}_{ki}) = \sqrt{(\mathbf{z}_i - \mathbf{m}_{ki})^{\mathrm{T}}\mathbf{S}_{ki}^{-1}(\mathbf{z}_i - \mathbf{m}_{ki})}$ is the Mahalanobis distance and J_i is the log PR of the ith segment.

2.4. *Thresholds selection*

With Eq. (13), we can derive the average J_i for each class (see Appendix A for details):

$$\begin{aligned} &E(J_i \,|\, k = L) \\ &\quad = \frac{1}{2}\left(\log\frac{|\mathbf{S}_{Ri}|}{|\mathbf{S}_{Li}|} + \mathrm{tr}((\mathbf{S}_{Li} + (\mathbf{m}_{Li} - \mathbf{m}_{Ri}) \cdot (\mathbf{m}_{Li} - \mathbf{m}_{Ri})^{\mathrm{T}}) \cdot \mathbf{S}_{Ri}^{-1} - \mathbf{I})\right) \\ &E(J_i \,|\, k = R) \\ &\quad = \frac{1}{2}\left(\log\frac{|\mathbf{S}_{Ri}|}{|\mathbf{S}_{Li}|} - \mathrm{tr}((\mathbf{S}_{Ri} + (\mathbf{m}_{Ri} - \mathbf{m}_{Li}) \cdot (\mathbf{m}_{Ri} - \mathbf{m}_{Li})^{\mathrm{T}}) \cdot \mathbf{S}_{Li}^{-1} - \mathbf{I})\right). \end{aligned} \tag{14}$$

According to SPRT,[21] for any given threshold pair η_L and η_R, the lower bound for the expectation of the accumulated segments number to make a correct decision, i.e. stopping time, for class k is $E(q \,|\, k)$. It is calculated by

$$\begin{aligned} &E(q \,|\, k = L) \\ &\quad = \inf_n \left(\sum_{i=1}^{n} E(J_i \,|\, k = L) \right. \\ &\quad \left. \geq (1 - p_{R|L}) \cdot \log\eta_L + p_{R|L} \cdot \log\eta_R\right) \\ &E(q \,|\, k = R) = \inf_n \left(\sum_{i=1}^{n} E(J_i \,|\, k = R)\right., \\ &\quad \leq \left(1 - \frac{1}{\eta_L}\right) \cdot \log\eta_R + \frac{1}{\eta_L} \cdot \log\eta_L\right) \end{aligned} \tag{15}$$

where $\inf_n(\cdot)$ is the minimum element of set of n. Combining with Eq. (12), Eq. (15) is then represented by

$$\begin{aligned} &E(q \,|\, k = L) = \inf_n \left(\sum_{i=1}^{n} E(J_i \,|\, k = L) \right. \\ &\quad \left. \geq (1 - \eta_R) \cdot \log\eta_L + \eta_R \cdot \log\eta_R\right) \\ &E(q \,|\, k = R) = \inf_n \left(\sum_{i=1}^{n} E(J_i \,|\, k = R) \right. \\ &\quad \left. \leq \left(1 - \frac{1}{\eta_L}\right) \cdot \log\eta_R + \frac{1}{\eta_L} \cdot \log\eta_L\right). \end{aligned} \tag{16}$$

Generally, $E(q \,|\, k = L)$ and $E(q \,|\, k = R)$ may be different. As the stopping time is a key point in the SA, we constrain the two thresholds by unifying $E(q \,|\, k)$ of two classes to be equal, that is, $E(q \,|\, k = L) = E(q \,|\, k = R) \triangleq q_E$. Then the thresholds to hit a certain stopping time (on average) are set by the following equation:

$$\begin{aligned} \xi_L &= \sum_{i=1}^{q_E} E(J_i \,|\, k = L) \\ \xi_R &= \sum_{i=1}^{q_E} E(J_i \,|\, k = R). \end{aligned} \tag{17}$$

For any given stopping time q_E, there is a corresponding threshold pair ξ_L and ξ_R. The decision rule with two thresholds ξ_L and ξ_R is

$$\begin{cases} \text{if } L_q \geq \xi_L, & \hat{k} = L \\ \text{if } L_q \leq \xi_R, & \hat{k} = R \\ \text{if } \xi_R < L_q < \xi_L, & \text{compute } L_{q+1}. \end{cases}$$

From this decision policy, we can see that other than assigning one of the two classes L and R, the decision functions may still be undecided and continue testing to the next observation. The "undecided" response keeps the number of errors (false positives or false negatives) low, which is useful for avoiding making excessive mistakes to speed up decisions, e.g. a BCI control wheelchair running into an obstacle.[13,17] In addition, we specify that when $q = q_E$, the decision rule is

$$\begin{cases} \text{if } L_q > 0, & \hat{k} = L \\ \text{if } L_q < 0, & \hat{k} = R. \end{cases}$$

Int. J. Neur. Syst. 2017.27. Downloaded from www.worldscientific.com
by ZHEJIANG UNIVERSITY on 11/29/17. For personal use only.

Till now, with the aforementioned decision rules, the consequent results, such as accuracy, mutual information (MI),[29,30] the steepness of MI[30] and average decision time, will only rely on the stopping time q_E. Depending on the actual specific needs, we could determine the optimal stopping time q_{opt} by maximizing certain evaluation parameter based on training data, such as accuracy. At the same time, the two thresholds are fixed.

3. Results

3.1. *Feature extraction*

In real application, the Morlet wavelets were digitized and limited with certain extension. As introduced in Section 2.2, the parameter optimization of the wavelet includes finding the modulation parameter K and the central frequency f. For a given f, K, the lengths of the wavelets should satisfy the e-folding requirement.[19] To account for this, the length of the Morlet wavelet corresponding to each rhythm was set to 1 s. Then a grid search is used to obtain the parameter vector that achieves the largest absolute value of t-statistics for each channel, based on all the segments of the training set. To find the optimal frequencies, the search range of f is expanded to $f_\mu = [6, 7, \ldots, 15]$ Hz and $f_\beta = [16, 17, \ldots, 30]$ Hz, respectively. K is within the range of $K = [5, 6, \ldots, 15]$. The optimal f, K are obtained when the absolute value of the t-statistics reaches the maximum.

Figure 1 shows the t-f-K searching surfaces of subjects S2003 in C3 and C4, and those of the other subjects are shown in Appendix B. The x, y and z axes of the searching surface correspond to f, K and t-statistics, respectively, in which t-statistics represents the separability of the two classes. From Fig. 1, the t-statistics corresponding to the wavelets of different central frequencies are quite different. The extreme points in Fig. 1 suggest the existence of optimal wavelet center frequencies, and the two peaks are located in the μ and β rhythms respectively. The frequency bands with the maximum separability show the individual variability. For subject S2003, the peak appears in μ rhythm. However, it appears in β rhythm for subject X11. Therefore, it is necessary to analyze the subject-specific frequency bands with an optimal Morlet wavelet method. Table 1 reveals the optimal parameters f, K for wavelet feature extraction of the four subjects in the first two datasets. It is clearly that some differences between C3 and C4 and between the subjects exist.

During the evidence accumulation, data points were fed to our method sequentially, segment by segment. To get the maximum time resolution, these segments have one sample overlap. So we have the average band power in the ERD/ERS time courses of EEG 4-D feature vector, as displayed in Fig. 2 and Appendix C. A clear hemispheric asymmetry could be observed in both subjects during imagination of one-sided hand movements. Moreover, specific ERD time courses were obtained for each subject. Starting

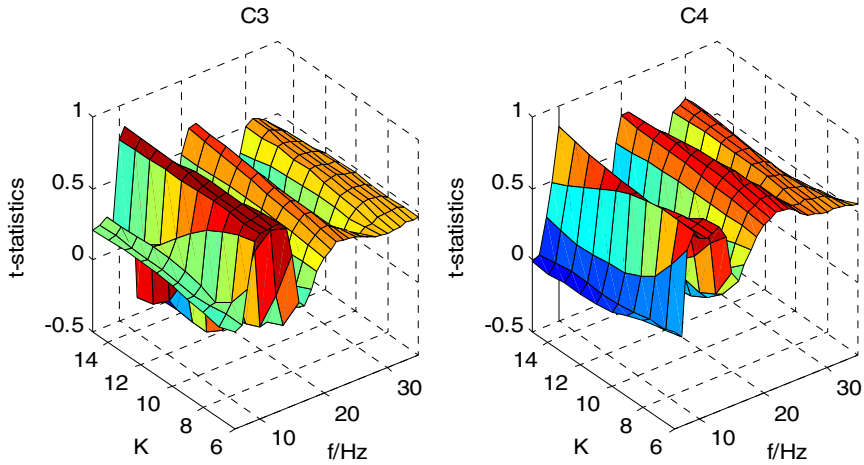


Fig. 1. The t-f-K searching surface in C3 and C4 of subject S2003. The x, y and z axes of the searching surface corresponds to and t-statistics, respectively.

Int. J. Neur. Syst. 2017.27. Downloaded from www.worldscientific.com
by ZHEJIANG UNIVERSITY on 11/29/17. For personal use only.

Table 1. Wavelet parameters for each subject.

Subject	Rhythm	f		K	
		C_3	C_4	C_3	C_4
S2003	μ	11	10	15	15
	β	22	22	15	10
O3	μ	11	12	6	9
	β	24	23	12	9
S4	μ	12	12	12	11
	β	22	23	12	8
X11	μ	12	14	12	11
	β	27	28	14	15

with cue presentation ($t = 3$ s), subject S2003 displayed a clear hemispheric differentiation in both μ and β band components. The power of the μ and β rhythms for the left-hand motor imagery displayed an amplitude enhancement (ERS) in electrode C3 and an amplitude attenuation (ERD) in electrode C4. During the right-hand motor task, the band power displayed amplitude attenuation (ERD) in electrode C3 and an amplitude enhancement (ERS) in electrode C4. In contrast with subject S2003, the reactivity patterns of μ and β rhythms for subject X11 were quite different. Although a bilateral desynchronization was found in μ band, there is little difference between the two conditions. The β rhythm first displayed a small bilateral ERD and then followed by a large ERS. The results of the average ERD/ERS time courses demonstrate that subject S2003 has larger distinction than subject X11 under the two tasks. The results shown in Figs. 1 and 2 suggest that the ERD/ERS varies between subjects. The results of the other subjects also show the variability of ERD/ERS from subject to subject.

3.2. *Time-accuracy tradeoff*

The result of SB developed by Lemm *et al.*[29] was demonstrated with the data of subject S2003. To compare the evidence accumulation process between BTSPRT and SB, we also use S2003 as an example subject. Due to the segments had one sample overlap, estimates of the classifier were updated every sample. The resulting time courses of the accumulative classification information, so-called accumulative evidence,[9] of the BTSPRT method and SB method are plotted in Fig. 3. From Fig. 3(a), we observe that the SB classification method gains information from around 4 s (3 s preparation period and 1 s window length). The cumulative Bayesian posterior probabilities reach the extreme at around 5.5 s, which indicates a peak decision confidence at that

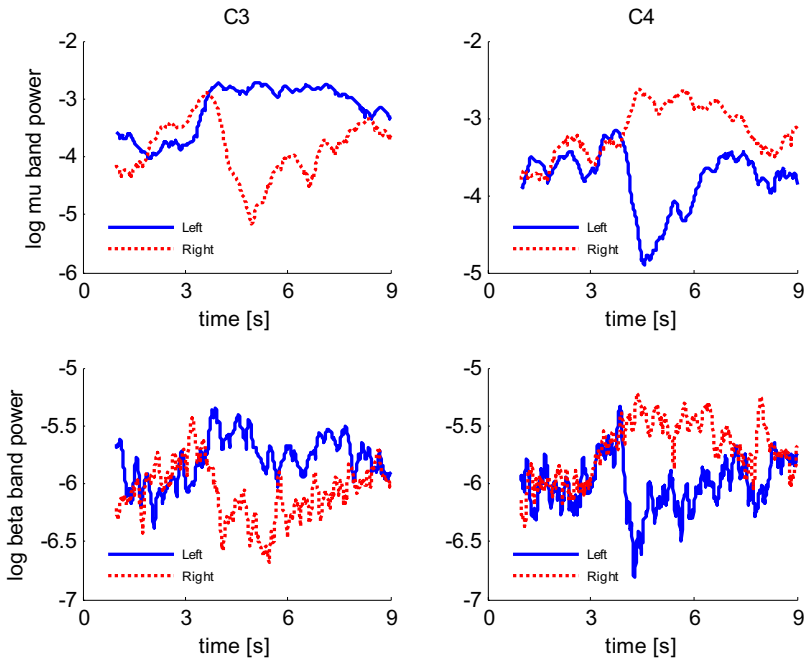
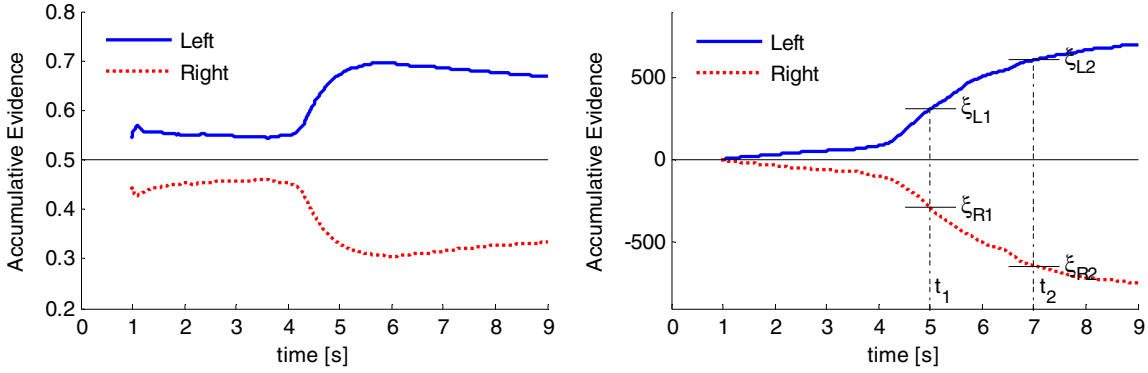


Fig. 2. The average ERD/ERS time courses of subject S2003 in the μ and β bands during imagined movement of the right (dot line) and left hand (solid line) for the C3 and C4 electrodes.



(a) (b)

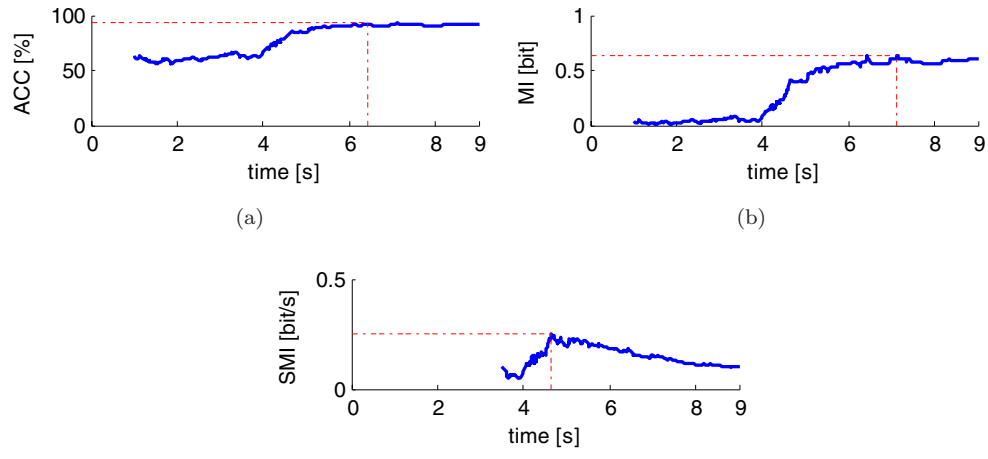
Fig. 3. The average accumulative process of classification information for subject S2003: (a) SB method and (b) BTSPRT method.

Int. J. Neur. Syst. 2017.27. Downloaded from www.worldscientific.com by ZHEJIANG UNIVERSITY on 11/29/17. For personal use only.

time. However, at the end of the trial, the accumulative information falls down. The result indicates that an effective control may take place in the middle of a trial. More evidence could not increase the classification performance any further.

Compared with the SB method, the cumulative evidence of BTSPRT is on the increase trend generally (shown in Fig. 3), which makes it possible to improve the accuracy with more evidence available. During the BTSPRT classification process, once the accumulative evidence exceeds one of the thresholds, an immediate decision will be given. Moreover, the thresholds can be adjusted to change the decision time. That is to say, with two broader threshold choices, a larger number of observations may be required to improve the accuracy and vice versa. This can be seen from Fig. 3(b), given the expected stopping time t_1, the thresholds are ξ_{L1}, ξ_{R1}. When the expected stopping time is set to be t_2, the thresholds change to be ξ_{L2}, ξ_{R2}. Obviously, it will achieve a higher accuracy with more decision time. This depicts the inherent tradeoff between decision time and accuracy of the BTSPRT method.

The time courses of the classification accuracy, the MI and steepness of MI (SMI, defined by $\mathrm{MI}(t)/(t - 3\,\mathrm{s})$ for $t > 3.5\,\mathrm{s}$) for subject S2003 are presented in Fig. 4. The SMI quantifies the response time. During the first 4 s, the classification performs at a rate no better than chance. Afterwards, there is a sharp increase in the classification accuracy, meanwhile reflecting a raising MI. The ascent of MI means increasing separation between the two motor



(a) (b) (c)

Fig. 4. The time course of the classification accuracy, the MI and the steepness of MI for subject S2003: (a) ACC; (b) MI; and (c) SMI.

imagery tasks. The maximum classification accuracy and MI are achieved at about 6.4 s and 7.1 s, respectively. The observation of nonmonotonic relationship between accuracy and the time is due to the limited data available, which leads to the final decision. With consideration about time, the maximum steepness of MI is obtained at around 4.6 s.

The general sequential framework of the present approach is customizable to suit different task objectives, such as improving accuracy, MI, or steepness of MI (SMI). The optimization of one particular objective will come at the expense of the others, in other words, the tradeoff between them.

3.3. *Classification results*

Two experiments were performed to evaluate the performance of BTSPRT. In the first one, for the purpose of benchmarking, we compared the classification accuracy among several state-of-the-art nonsequential classifiers used in the BCI community.[31–33] These methods were applied to the two BCI competition datasets described in Sec. 2.1 with features based on the optimal wavelet method. Those methods were particle swarm optimization-based radial basis function network (PSO-RBFN),[31] improved fuzzy support vector classifier (IFFSVC),[31] fuzzy inference system (FIS),[32] support vector machine (SVM),[32] Gaussian process (GP),[33] linear discriminant analysis (LDA)[34] and the adaptive linear LDA (ALDA).[35,36] As mentioned before, the optimal stopping time q_{opt} is determined by a certain goal. When the goal is set to achieve high classification accuracy, q_{opt} is determined by maximizing the accuracy of the training data. Then, the two thresholds ξ_L and ξ_R are set by Eq. (17). The results of all the methods on the competition test data are shown in Table 1, in which "Avg." means the averaged accuracy of all the subjects. By applying the obtained effective features with the optimal wavelet method on the SPRT, a classification accuracy of 85.9% is achieved. They are greater than that achieved by the methods 4–8. Overall, the classification accuracy is higher when the proposed method is adopted. In terms of average accuracy, BTSPRT reached similar results as the PSO-RBFN in Ref. 31 and outperforms other classifiers. Nevertheless, Table 2 also shows that the BTSPRT may not obtain the best results for all subjects due to subject-to-subject differences.

Table 2. Accuracy (%) of different classifiers on the BCI competition datasets.

Num.	Method	S2003	O3	S4	X11	Avg.
1	BTSPRT	90.7	88.3	79.4	85.0	85.9
2	PSO-RBFN	86.7	95.5	79.9	78.9	85.3
3	IFFSVC	85.6	95.1	75.8	78.8	83.3
4	FIS	82.9	86.7	74.7	75.7	80.0
5	SVM	83.5	86.8	75.9	75.4	80.4
6	GP	82.1	89.3	72.9	74.3	79.7
7	LDA	77.1	92.9	81.5	80.4	83.0
8	ALDA	79.2	78.6	76.3	77.0	77.8

After analyzing the overall accuracy results, it can be concluded that BTSPRT outperforms the state-of-the-art classification methods on the standard datasets.

Additionally, MI criterion is often used to compare the performance of different methods. It is also important to consider the stopping time for speedy BCI applications without sacrificing accuracy.[24] Therefore, in the second test, we compared BTSPRT versus the competition winner, the SB posterior probability method,[19] on the two BCI competition datasets and one online experiment in terms of the achieved classification accuracy, the MI and stopping time. The results are provided in Table 3. It should be noted that the stopping time T in Table 3 is the time period from the start of motor imagery to the decision completion.

From it, we can see that the SB method has a 79.2% average accuracy, whereas that of the BTSPRT is increased to 83.4% with the same optimal wavelet-based features. Moreover, we observed that the BTSPRT classifier outperforms the SB classifier in terms of classification accuracy across all the subjects except for subject S4. Moreover, the average stopping time is shortened with BTSPRT. In every case, BTSPRT gives better accuracy and yields a faster decision than SB does. The MI for BTSPRT was consequently higher than that of SB. Furthermore, a paired t-test analysis was used to compare the classification accuracies of the two methods. The paired t-test result confirms that with SPRT the accuracies across subjects are significantly higher than that with SB ($t = 7.65 > t_{0.01}(13) = 1.77$). The statistical results on the MI also indicate that the BTSPRT has a significantly better performance than that of the SB ($t = 8.14 > t_{0.01}(13) = 1.77$).

Int. J. Neur. Syst. 2017.27. Downloaded from www.worldscientific.com
by ZHEJIANG UNIVERSITY on 11/29/17. For personal use only.

Table 3. Performance comparison between BTSPRT and SB methods with the three datasets.

Method	SB			BTSPRT		
Subject	ACC (%)	MI (bit)	T (s)	ACC (%)	MI (bit)	T (s)
S2003	88.6	0.488	3.18	90.7	0.554	2.16
O3	84.4	0.375	2.48	88.3	0.479	1.53
S4	78.3	0.245	3.74	79.5	0.268	3.00
X11	83.7	0.359	3.66	85.0	0.390	2.85
L1	80.0	0.278	3.04	82.0	0.320	3.28
L2	80.0	0.278	2.80	86.0	0.416	2.99
L3	76.0	0.205	3.01	82.0	0.320	2.78
L4	78.0	0.240	2.53	82.0	0.320	2.81
L5	72.0	0.145	2.89	78.0	0.240	2.56
L6	74.0	0.173	3.27	78.0	0.240	3.11
L7	78.0	0.240	2.76	82.0	0.320	2.72
L8	82.0	0.320	2.74	86.0	0.416	3.27
L9	78.0	0.240	2.99	84.0	0.366	3.05
L10	76.0	0.205	2.63	84.0	0.366	2.64
Avg.±	79.2	0.271	3.01	83.4	0.358	2.77
Std	4.38	0.090	0.367	3.67	0.089	0.463

Note: *ACC, accuracy; MI, mutual information; T, time.

Int. J. Neur. Syst. 2017.27. Downloaded from www.worldscientific.com
by ZHEJIANG UNIVERSITY on 11/29/17. For personal use only.

The average stopping time is decreased from 3.01 s with SB to 2.77 s with BTSPRT, although there is no significant difference between them ($t = 1.61 < t_{0.01}(13) = 1.77$). Our results are consistent with the literature results that show that the BTSPRT method could achieve better classification performance than the SB does with even fewer test items.[37]

4. Conclusion

The BCI performance in a specific application can be expressed as speed or accuracy or both. If the goal involves transmission of relatively simple commands, for example, to answer a yes/no question, change room temperature, select a menu item, or control hand grasp through a neural prosthesis, the accuracy is often more important than speed. For time-sensitive BCI applications, such as controlling a motorized wheelchair, the control speed can be emphasized over accuracy in performance. In addition, in BCI gaming applications, a tradeoff between timing precision and classification precision has to be found.

For this aim, we present a balanced threshold SPRT method in conjunction with optimal wavelet analysis method to recognize mental states. A user-specific method was first developed to determine features based on a grid search on the wavelet parameters. With the accumulative evidence curve, the proposed BTSPRT set the two constrained thresholds based on a desired expected stopping time. The BTSPRT method adds the benefit of a customizable tradeoff between accuracy and decision speed. Specifically, the thresholds in this method were determined without predefined error probabilities. Using standardized datasets and online test, improved success rates and decision times were demonstrated.

Although this study suggests that BTSPRT would be useful to balance the speed and accuracy for different BCI applications, we realized that further work requires investigation. Future work will attempt to validate the proposed method with a larger dataset and to implement it in our BCI-actuated robotic system.[38] Moreover, we will investigate the multiway SPRT theory and its application to the multi-class BCI systems.

Acknowledgments

This work was supported in part by the grants from the Natural Science Foundation of China (NSFC) (61005088 and 61573079), the Open Research Foundation of State Key Lab. of Digital Manufacturing

Equipment and Technology in Huazhong University (DMETKF2015007) and the Fundamental Research Funds for the Central Universities to Dr Liu and by the U.S. National Science Foundation Cyber Enabled Discovery and Innovation (CDI) program grant to Dr Thakor, numbers ECCS 0835632 and ECCS 0835554. The authors are also very grateful to the technical editor and anonymous reviewers for their valuable and constructive comments.

Int. J. Neur. Syst. 2017.27. Downloaded from www.worldscientific.com by ZHEJIANG UNIVERSITY on 11/29/17. For personal use only.

Appendix A. Proof of Eq. (14)

For convenience, we omit the subscript i for each segment. Then the log probability ratio of each segment J is

$$J = \frac{1}{2}\left(\log\frac{|\mathbf{S}_R|}{|\mathbf{S}_L|} + D_M^2(\mathbf{z},\mathbf{m}_R) - D_M^2(\mathbf{z},\mathbf{m}_L)\right)$$
$$= \frac{1}{2}\begin{pmatrix}\log\dfrac{|\mathbf{S}_R|}{|\mathbf{S}_L|} + (\mathbf{z}-\mathbf{m}_R)^{\mathrm{T}}\mathbf{S}_R^{-1}(\mathbf{z}-\mathbf{m}_R)\\ -(\mathbf{z}-\mathbf{m}_L)^{\mathrm{T}}\mathbf{S}_L^{-1}(\mathbf{z}-\mathbf{m}_L)\end{pmatrix}, \tag{A.1}$$

$$J = \frac{1}{2}\left(\log\frac{|\mathbf{S}_R|}{|\mathbf{S}_L|} + D_M^2(\mathbf{z},\mathbf{m}_R) - D_M^2(\mathbf{z},\mathbf{m}_L)\right)$$
$$= \frac{1}{2}\begin{pmatrix}\log\dfrac{|\mathbf{S}_R|}{|\mathbf{S}_L|} + (\mathbf{z}-\mathbf{m}_R)^{\mathrm{T}}\mathbf{S}_R^{-1}(\mathbf{z}-\mathbf{m}_R)\\ -(\mathbf{z}-\mathbf{m}_L)^{\mathrm{T}}\mathbf{S}_L^{-1}(\mathbf{z}-\mathbf{m}_L)\end{pmatrix}. \tag{A.2}$$

When the class is L, we can write the average J as

$$E(J\,|\,k=L)$$
$$= E\left(\frac{1}{2}\begin{pmatrix}\log\dfrac{|\mathbf{S}_R|}{|\mathbf{S}_L|} + (\mathbf{z}_L-\mathbf{m}_R)^{\mathrm{T}}\\ \cdot\,\mathbf{S}_R^{-1}(\mathbf{z}_L-\mathbf{m}_R)\\ -(\mathbf{z}_L-\mathbf{m}_L)^{\mathrm{T}}\mathbf{S}_L^{-1}(\mathbf{z}_L-\mathbf{m}_L)\end{pmatrix}\right)$$
$$= \frac{1}{2}\begin{pmatrix}\log\dfrac{|\mathbf{S}_R|}{|\mathbf{S}_L|} + E((\mathbf{z}_L-\mathbf{m}_R)^{\mathrm{T}}\\ \cdot\,\mathbf{S}_R^{-1}(\mathbf{z}_L-\mathbf{m}_R))\\ -E\,(\mathbf{z}_L-\mathbf{m}_L)^{\mathrm{T}}\mathbf{S}_L^{-1}(\mathbf{z}_L-\mathbf{m}_L)\end{pmatrix}. \tag{A.3}$$

The following facts are used throughout:

$$E(\mathbf{x}^{\mathrm{T}}\mathbf{x}) = E(\mathrm{tr}(\mathbf{x}^{\mathrm{T}}\mathbf{x})) = E(\mathrm{tr}(\mathbf{x}\mathbf{x}^{\mathrm{T}}))$$
$$= \mathrm{tr}(E(\mathbf{x}\mathbf{x}^{\mathrm{T}})), \tag{A.4}$$

in which $\mathbf{x} \in R^{m\times 1}$.

$$\mathrm{tr}(\mathbf{A}\mathbf{B}) = \mathrm{tr}(\mathbf{B}\mathbf{A}), \tag{A.5}$$

in which $\mathbf{A} \in R^{m\times n}, \mathbf{B} \in R^{n\times m}$.

Therefore,

$$\begin{aligned}&E((\mathbf{z}_L-\mathbf{m}_L)^{\mathrm{T}}\mathbf{S}_L^{-1}(\mathbf{z}_L-\mathbf{m}_L))\\ &\quad= E(\mathrm{tr}((\mathbf{z}_L-\mathbf{m}_L)^{\mathrm{T}}\mathbf{S}_L^{-1}(\mathbf{z}_L-\mathbf{m}_L)))\\ &\quad= E(\mathrm{tr}((\mathbf{z}_L-\mathbf{m}_L)(\mathbf{z}_L-\mathbf{m}_L)^{\mathrm{T}}\mathbf{S}_L^{-1}))\\ &\quad= \mathrm{tr}(E((\mathbf{z}_L-\mathbf{m}_L)(\mathbf{z}_L-\mathbf{m}_L)^{\mathrm{T}}\mathbf{S}_L^{-1}))\\ &\quad= \mathrm{tr}(E((\mathbf{z}_L-\mathbf{m}_L)(\mathbf{z}_L-\mathbf{m}_L)^{\mathrm{T}})\mathbf{S}_L^{-1})\\ &\quad= \mathrm{tr}(\mathbf{S}_L\mathbf{S}_L^{-1})\\ &\quad= \mathrm{tr}(\mathbf{I})\end{aligned} \tag{A.6}$$

and

$$\begin{aligned}&E((\mathbf{z}_L-\mathbf{m}_R)^{\mathrm{T}}\mathbf{S}_R^{-1}(\mathbf{z}_L-\mathbf{m}_R))\\ &\quad= E(((\mathbf{z}_L-\mathbf{m}_L)+(\mathbf{m}_L-\mathbf{m}_R))^{\mathrm{T}}\\ &\qquad\cdot\,\mathbf{S}_R^{-1}((\mathbf{z}_L-\mathbf{m}_L)+(\mathbf{m}_L-\mathbf{m}_R)))\\ &\quad= E((\mathbf{z}_L-\mathbf{m}_L)^{\mathrm{T}}\mathbf{S}_R^{-1}(\mathbf{z}_L-\mathbf{m}_L))\\ &\qquad+E((\mathbf{z}_L-\mathbf{m}_L)^{\mathrm{T}}\mathbf{S}_R^{-1}(\mathbf{m}_L-\mathbf{m}_R))\\ &\qquad+E((\mathbf{m}_L-\mathbf{m}_R)^{\mathrm{T}}\mathbf{S}_R^{-1}(\mathbf{z}_L-\mathbf{m}_L))\\ &\qquad+E((\mathbf{m}_L-\mathbf{m}_R)^{\mathrm{T}}\mathbf{S}_R^{-1}(\mathbf{m}_L-\mathbf{m}_R))\\ &\quad= E((\mathbf{z}_L-\mathbf{m}_L)^{\mathrm{T}}\mathbf{S}_R^{-1}(\mathbf{z}_L-\mathbf{m}_L))\\ &\qquad+\mathbf{0}+\mathbf{0}+(\mathbf{m}_L-\mathbf{m}_R)^{\mathrm{T}}\mathbf{S}_R^{-1}(\mathbf{m}_L-\mathbf{m}_R)\\ &\quad= E(\mathrm{tr}((\mathbf{z}_L-\mathbf{m}_L)(\mathbf{z}_L-\mathbf{m}_L)^{\mathrm{T}}\mathbf{S}_R^{-1}))\\ &\qquad+\mathrm{tr}((\mathbf{m}_L-\mathbf{m}_R)(\mathbf{m}_L-\mathbf{m}_R)^{\mathrm{T}}\mathbf{S}_R^{-1})\\ &\quad= \mathrm{tr}(E((\mathbf{z}_L-\mathbf{m}_L)(\mathbf{z}_L-\mathbf{m}_L)^{\mathrm{T}})\mathbf{S}_R^{-1})\\ &\qquad+\mathrm{tr}((\mathbf{m}_L-\mathbf{m}_R)(\mathbf{m}_L-\mathbf{m}_R)^{\mathrm{T}}\mathbf{S}_R^{-1})\\ &\quad= \mathrm{tr}(\mathbf{S}_L\mathbf{S}_R^{-1})+\mathrm{tr}((\mathbf{m}_L-\mathbf{m}_R)\\ &\qquad\cdot(\mathbf{m}_L-\mathbf{m}_R)^{\mathrm{T}}\mathbf{S}_R^{-1}).\end{aligned} \tag{A.7}$$

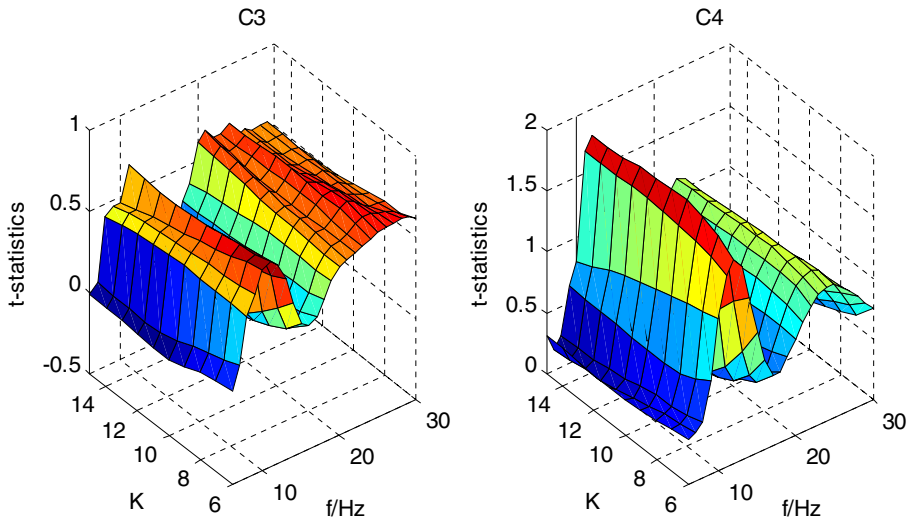
Using Eqs. (A.4), (A.6) and (A.7) we can show by substitution that (A.3), we obtain

$$E(J_i\,|\,k=L)$$
$$= \frac{1}{2}\begin{pmatrix}\log\dfrac{|\mathbf{S}_{Ri}|}{|\mathbf{S}_{Li}|} + \mathrm{tr}((\mathbf{S}_{Li}+(\mathbf{m}_{Li}-\mathbf{m}_{Ri})\\ \cdot\,(\mathbf{m}_{Li}-\mathbf{m}_{Ri})^{\mathrm{T}})\cdot\mathbf{S}_{Ri}^{-1}-\mathbf{I})\end{pmatrix}. \tag{A.8}$$

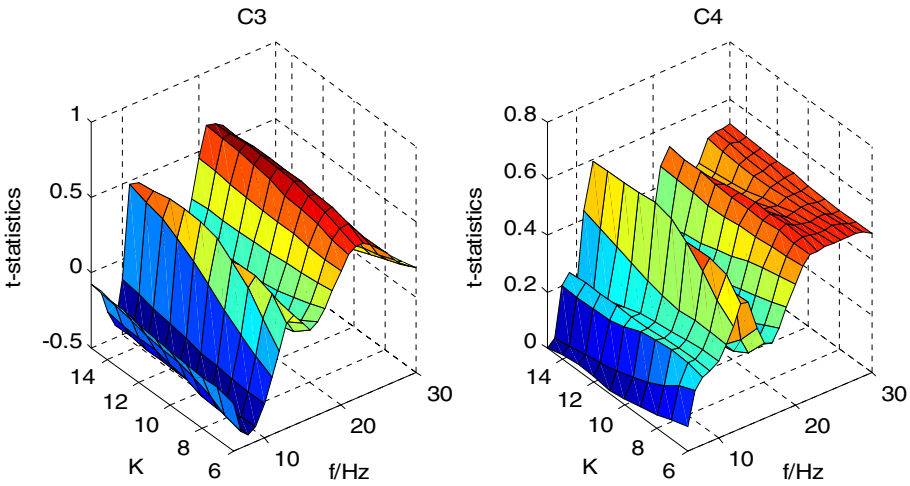
When the class is R, similar derivation leads to the conclusion that

$$E(J_i \mid k = R) = \frac{1}{2}\begin{pmatrix} \log\dfrac{|\mathbf{S}_{Ri}|}{|\mathbf{S}_{Li}|} - \mathrm{tr}((\mathbf{S}_{Ri} + (\mathbf{m}_{Ri} - \mathbf{m}_{Li}) \\ \cdot (\mathbf{m}_{Ri} - \mathbf{m}_{Li})^{\mathrm{T}}) \cdot \mathbf{S}_{Li}^{-1} - \mathbf{I}) \end{pmatrix}. \tag{A.9}$$

Appendix B. The *t-f-K* Searching Surfaces in C3 and C4

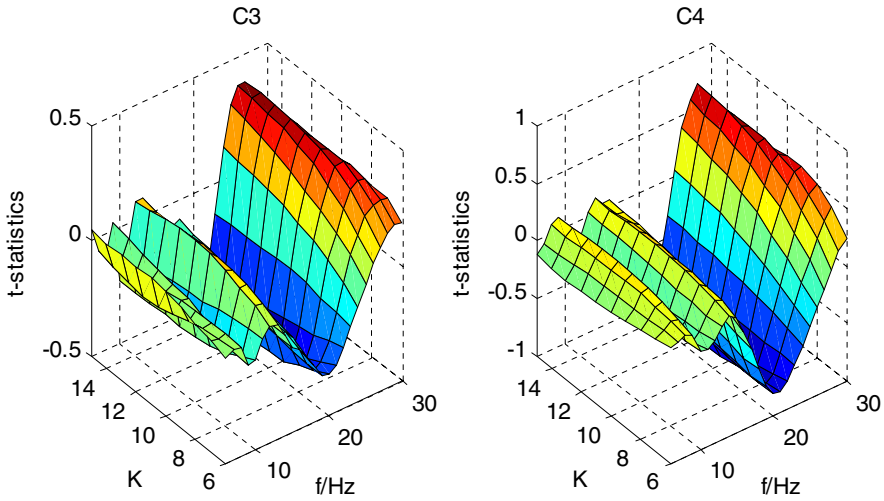


(B.1) Subject O3



(B.2) Subject S4

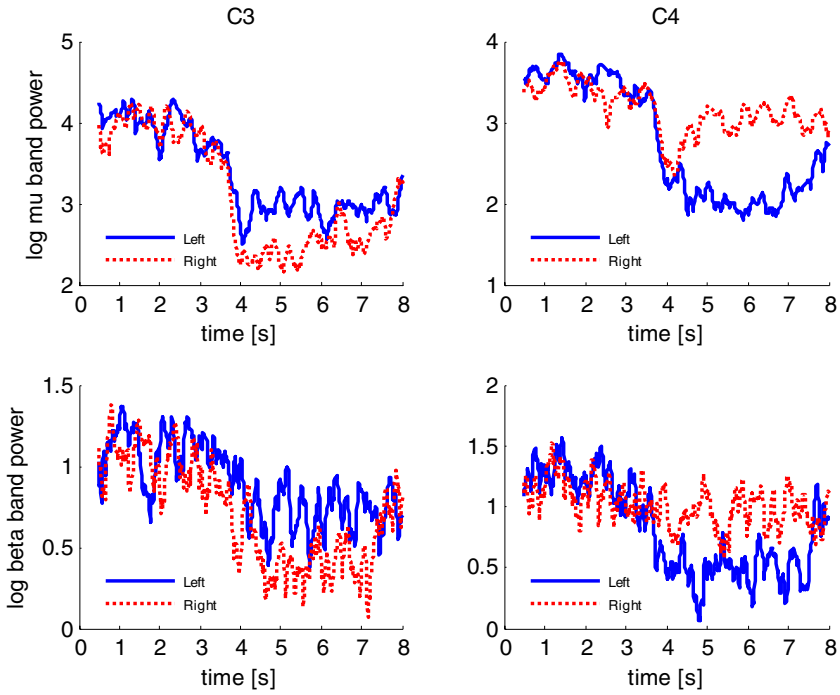
Int. J. Neur. Syst. 2017.27. Downloaded from www.worldscientific.com
by ZHEJIANG UNIVERSITY on 11/29/17. For personal use only.



(B.3) Subject X11

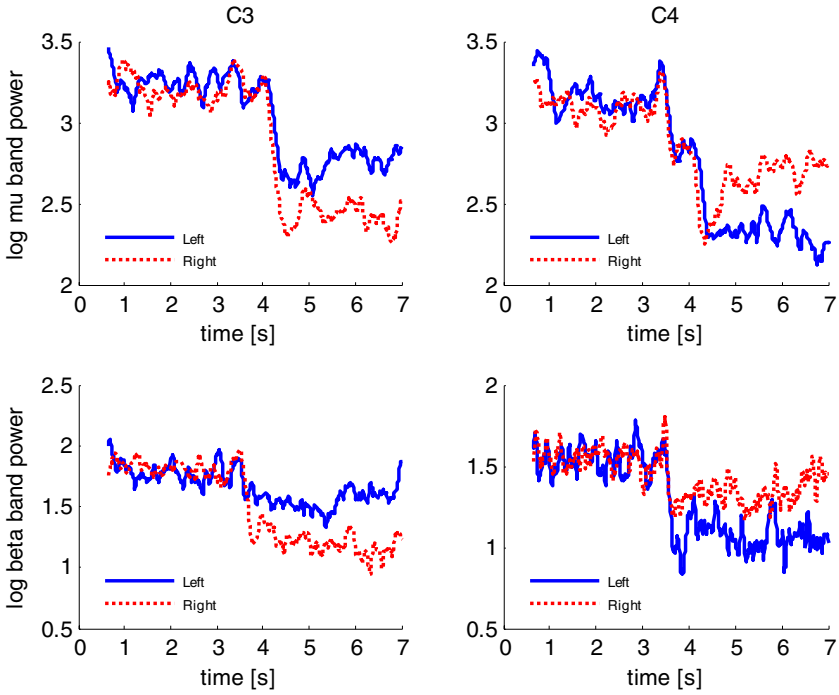
Int. J. Neur. Syst. 2017.27. Downloaded from www.worldscientific.com
by ZHEJIANG UNIVERSITY on 11/29/17. For personal use only.

Appendix C. The Average ERD/ERS Time Courses of Other Subjects in the μ and β Bands

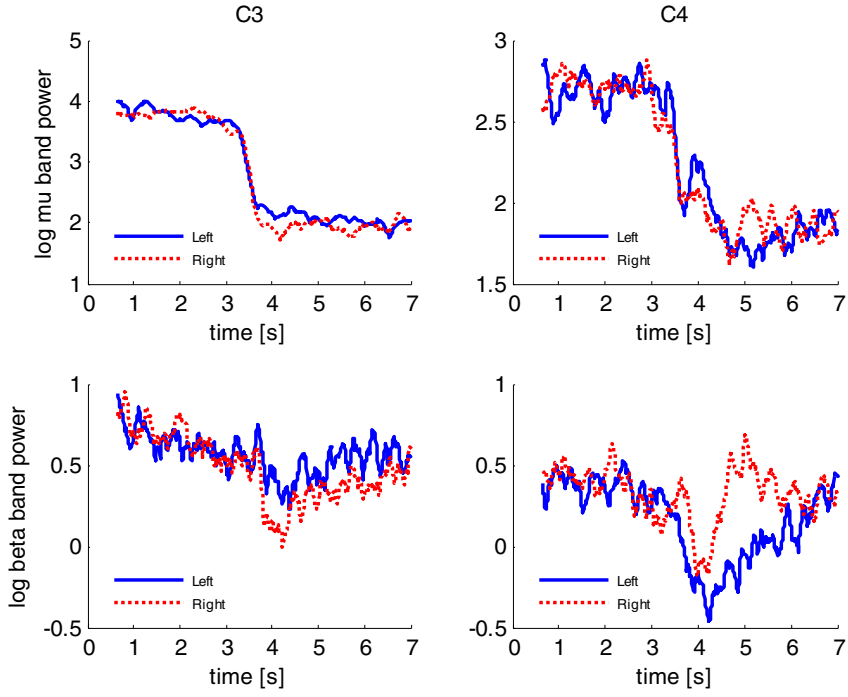


(C.1) Subject O3

Int. J. Neur. Syst. 2017.27. Downloaded from www.worldscientific.com
by ZHEJIANG UNIVERSITY on 11/29/17. For personal use only.



(C.2) Subject S4



(C.3) Subject X11

Int. J. Neur. Syst. 2017.27. Downloaded from www.worldscientific.com
by ZHEJIANG UNIVERSITY on 11/29/17. For personal use only.

References

1. J. Li, H. Ji, L. Cao, D. Zang, R. Gu, B. Xia and Q. Wu, Evaluation and application of a hybrid brain computer interface for real wheelchair control with multi-degrees of freedom, *Int. J. Neural Syst.* **24**(4) (2014) 1450014.
2. J. Li, J. Liang, Q. Zhao, K. Hong and L. Zhang, Design of assistive wheelchair system directly steered by human thoughts, *Int. J. Neural Syst.* **23**(3) (2013) 1350013.
3. J. Jin, E. W. Sellers, S. J. Zhou, Y. Zhang, X. Y. Wang and A. Cichocki, A P300 Brain–computer interface based on a modification of the mismatch negativity paradigm, *Int. J. Neural Syst.* **25**(3) (2015) 1550011.
4. A. Burns, H. Adeli and J. A. Buford, Brain–computer interface after nervous system injury, *Neuroscientist* **20**(6) (2014) 639–651.
5. W. Y. Hsu, Continuous EEG signal analysis for asynchronous BCI application, *Int. J. Neural Syst.* **21**(4) (2011) 335–350.
6. F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche and B. Arnaldi, A review of classification algorithms for EEG-based brain–computer interfaces, *J. Neural Eng.* **4**(2) (2007) R1–R13.
7. M. Ahmadlou and H. Adeli, Enhanced probabilistic neural network with local decision circles: A robust classifier, *Integr. Comput-Aided Eng.* **17**(3) (2010) 197–210.
8. M. N. Shadlen and R. Kiani, Decision making as a window on cognition, *Neuron* **80**(3) (2013) 791–806.
9. J. I. Gold and M. N. Shadlen, The neural basis of decision making, *Annu. Rev. Neurosci.* **30** (2007) 535–574.
10. H. R. Heekeren, S. Marrett and L. G. Ungerleider, The neural systems that mediate human perceptual decision making, *Nat. Rev. Neurosci.* **9**(6) (2008) 467–479.
11. P. Domenech and J. Dreher, Decision threshold modulation in the human brain, *J. Neurosci.* **30**(43) (2010) 14305–14317.
12. A. Ortiz-Rosario and H. Adeli, Brain–computer interface technologies: From signal to action, *Rev. Neurosci.* **24**(5) (2013) 537–552.
13. J. D. R. Millán, Brain-controlled robots, *IEEE Intell. Syst.* **23**(3) (2008) 74–76.
14. G. Krausz, R. Scherer, G. Korisek and G. Pfurtscheller, Critical decision-speed and information transfer in the "Graz Brain–Computer Interface", *Appl. Psychophysiol. Biofeedb.* **28**(3) (2003) 233–240.
15. R. Liu, Y. X. Wang and L. Zhang, An FDES-based shared control method for asynchronous brain-actuated robot, *IEEE Trans. Cybern.* **46**(6) (2016) 1452–1462.
16. Y. Zhang, G. Zhou, J. Jin, Q. Zhao, X. Wang and A. Cichocki, Aggregation of sparse linear discriminant analyses for event-related potential classification in brain–computer interface, *Int. J. Neural Syst.* **24**(1) (2014) 1450003.
17. R. Liu, G. Newman, S. H. Ying and N. V. Thakor, Improved BCI performance with sequential hypothesis testing, in *Proc. 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Boston, USA (2011), pp. 4215–4218.
18. R. Scherer, F. Lee, A. Schlögl, R. Leeb, H. Bischof and G. Pfurtscheller, Toward self-paced brain–computer communication: Navigation through virtual worlds, *IEEE Trans. Biomed. Eng.* **55**(2 Pt 1) (2008) 675–682.
19. S. Lemm, C. Schäfer and G. Curio, BCI competition 2003-data set III: Probabilistic modeling of sensorimotor μ rhythms for classification of imaginary hand movements, *IEEE Trans. Biomed. Eng.* **51**(6) (2004) 1077–1080.
20. G. Gangadhar, R. Chavarriaga and J. D. R. Millán, Fast recognition of anticipation-related potentials, *IEEE Trans. Biomed. Eng.* **56**(4) (2009) 1257–1260.
21. A. Wald (ed.), *Sequential Analysis* (Dover Publications, New York, 2004).
22. C. Anderson, E. Forney, D. Hains and A. Natarajan, Reliable identification of mental tasks using time-embedded EEG and sequential evidence accumulation, *J. Neural Eng.* **8**(2) (2011) 025023.
23. H. G. Faulkner, A. Myrden, M. Li and K. Mamun, Sequential hypothesis testing for automatic detection of task-related changes in cerebral perfusion in a brain–computer interface, *Neurosci. Res.* **100** (2015) 29–38.
24. B. Blankertz, K. R. Müller, T. V. G. Curio, G. Schalk, J. Wolpaw, A. Schlögl, C. Neuper, G. Pfurtscheller, T. Hinterberger and M. S. N. Birbaumer, The BCI competition 2003: Progress and perspectives in detection and discrimination of EEG single trials, *IEEE Trans. Biomed. Eng.* **51**(6) (2004) 1044–1051.
25. B. Blankertz, K. R. Müller, D. J. Krusienski, G. Schalk, J. R. Wolpaw, A. Schlögl, G. Pfurtscheller, J. D. R. Millán, M. Schroder and N. Birbaumer, The BCI competition III: Validating alternative approaches to actual BCI problems, *IEEE Trans. Neural Syst. Rehabil.* **14** (2006) 153–159.
26. G. R. Müller-Purz, C. Pokorny, D. S. Klobassa and P. Horki, A single-switch BCI based and passive and imagined movements: Toward restoring communication in minimally conscious patients, *Int. J. Neural. Syst.* **23**(2) (2013) 1250037.
27. H. Adeli, Z. Zhou and N. Dadmehr, Analysis of EEG records in an epileptic patient using wavelet transform, *J. Neurosci. Methods* **12**(1) (2003) 69–87.

28. H. Adeli, S. Ghosh-Dastidar and N. Dadmehr, A wavelet-chaos methodology for analysis of EEGs and EEG sub-bands to detect seizure and epilepsy, *IEEE Trans. Biomed. Eng.* **54**(2) (2007) 205–211.
29. S. Lemm, C. Schafer and G. Curio, Aggregating classification accuracy across time: Application to single trial EEG, *Adv. Neural Inform. Process. Syst.* **19** (2007) 825–832.
30. A. Schlögl, J. Kronegg, J. E. Huggins and S. G. Mason, Evaluation criteria for BCI research, in *Toward Brain–Computer Interfacing* (MIT Press, Cambridge, 2007), pp. 327–342.
31. E. Cinar and F. Sahin, New classification techniques for electroencephalogram (EEG) signals and a real-time EEG control of a robot, *Neural Comput. Appl.* **22**(1) (2013) 29–39.
32. F. Lotte, The use of fuzzy inference system for classification in EEG-based brain computer interfaces, in *Proc. 3rd International Brain–Computer Interface Workshop and Training Course*, Graz, Austria (2006), pp. 12–13.
33. M. J. Zhong, F. Lotte and M. Girolami, Classifying EEG for brain computer interfaces using Gaussian processes, *Pattern Recogn. Lett.* **29**(3) (2008) 354–359.
34. N. Brodu, F. Lotte and A. Lécuyer, Comparative study of band-power extraction techniques for motor imagery classification, *IEEE Symp. Computational Intelligence*, *Cognitive Algorithms*, *Mind*, *and Brain* (*CCMB*), Paris (2011), pp. 1–6.
35. C. Vidaurre, A. Schlögl, R. Cabeza, R. Scherer and G. Pfurtscheller, Study of on-line adaptive discriminant analysis for EEG-based brain computer interfaces, *IEEE Trans. Biomed. Eng.* **54**(3) (2007) 550–556.
36. W. Y. Hsu, EEG-based motor imagery classification using enhanced active segment selection and adaptive classifier, *Comp. Biol. Med.* **41**(8) (2011) 633–639.
37. J. A. Spray and M. D. Reckase, Comparison of SPRT and sequential Bayes procedures for classifying examinees into two categories using a computerized test, *J. Educ. Behav. Stat.* **21**(4) (1996) 405–414.
38. R. Liu, Y. X. Wang and L. Zhang, An FDES-based shared control method for asynchronous brain-actuated robot, *IEEE Trans. Cyber.* **46**(6) (2016) 1452–1462.
39. C. Guger, B. Allison, B. Gro Großwindhager, R. Prückl, C. Hintermüller, C. Kapeller, M. Bruckner, G. Krausz and G. Edlinger, How many people could use an SSVEP BCI? *Front. Neurosci.* **6**(6) (2012) 169, doi:10.3389/fnins.2012.00169.

Int. J. Neur. Syst. 2017.27. Downloaded from www.worldscientific.com
by ZHEJIANG UNIVERSITY on 11/29/17. For personal use only.