

# Single-Trial Event-Related Potential Extraction Through One-unit ICA-with-Reference

Wee Lih Lee<sup>1</sup>, Tele Tan<sup>2</sup>, Torbjörn Falkmer<sup>3 4 5</sup> and Yee Hong Leung<sup>1</sup>

<sup>1</sup> Department of Electrical and Computer Engineering, Curtin University, Perth, Australia

<sup>2</sup> Department of Mechanical Engineering, Curtin University, Perth, Australia

<sup>3</sup> School of Occupational Therapy and Social Works, Curtin University, Perth, Australia

<sup>4</sup> Rehabilitation Medicine, Department of Medicine and Health Sciences (IMH), Faculty of Health Sciences, Linköping University

<sup>5</sup> Pain and Rehabilitation Centre, UHL, County Council, Linköping, Sweden

E-mail: T.Tan@curtin.edu.au

**Abstract.** *Objective:* In recent years, ICA has been one of the more popular methods for extracting event-related potential (ERP) at the single-trial level. It is a blind source separation technique that allows the extraction of an ERP without making strong assumptions on the temporal and spatial characteristics of an ERP. However, the problem with traditional ICA is that the extraction is not direct and is time-consuming due to the need for source selection processing. In this paper, the application of an one-unit ICA-with-Reference (ICA-R), a constrained ICA method, is proposed. *Approach:* In cases where the time-region of the desired ERP is known *a priori*, this time information is utilized to generate a reference signal, which is then used for guiding the one-unit ICA-R to extract the source signal of the desired ERP directly. *Main Results:* Our results showed that, as compared to traditional ICA, ICA-R is a more effective method for analysing ERP because it avoids manual source selection and it requires less computation thus resulting in faster ERP extraction. *Significance:* In addition to that, since the method is automated, it reduces the risks of any subjective bias in the ERP analysis. It is also a potential tool for extracting the ERP in online application.

**Keywords:** Constrained independent component analysis, Event-related potential, Independent component analysis, Independent component analysis with reference, P300, single-trial extraction

Submitted to: *J. Neural Eng.*

## 1. Introduction

Event related potentials (ERPs) are electrical activities in the brain that occur milliseconds to a few hundred milliseconds after a specific internal or external event occurs [1, 2]. They index the chronological neural processes that are responsible for how a human senses, perceives, recognizes and reacts towards an event. In electroencephalography (EEG), ERPs often appear as a series of positive and negative peaks each with a distinct scalp distribution and specific range of occurrence time following stimulus onset. Since each ERP reflects different mental states within the brain, they are useful in studying the underlying neurocognitive mechanisms and provide a means of assessing different neurological disorders [3].

The traditional method for analyzing ERPs usually begins with segmenting an EEG recording into trials/epochs according to the stimulus onset. Since ERPs have relatively small amplitudes compared to the on-going background EEG activities and noise artifacts, it is common to apply a trial-averaging technique to suppress the random signals that are not time-locked across trials. As such, the signals can be smoothed and the overall representation of the ERPs can be obtained [3, 4, 5, 6]. However, there are a few drawbacks with the aforementioned trial-averaging technique. Firstly, trial-averaging assumes the ERP waveforms are invariant across trials. This assumption however, is not valid because numerous findings have suggested that ERPs can vary greatly even under the same stimulus condition [7, 8]. For example, a person's physiological conditions, such as mental fatigue, habituation and attentiveness can contribute towards trial variations in the ERPs [9]. Secondly, the trial-averaging technique inhibits trial-to-trial analysis where important information, which may reflect the changing pattern of ERP amplitude and latency across trials, is lost due to the averaging process.

Since the time course of the ERP often provides a richer and more valuable picture of the neural activity, the ability to examine the dynamics of ERPs at single-trial level is always desired. For this reason, various approaches have been proposed for robust single-trial ERP extraction. These techniques can be divided into two main types: single-channel and multi-channel ERP extraction. Single-channel methods usually involve denoising and decomposition of a temporal signal from a selected EEG channel. For example, wavelet de-

noising methods have been introduced in previous research [10, 11] to remove the spontaneous EEG allowing the desired ERP to be reliably estimated. A framework which treats each EEG epoch as a linear combination of multiple ERP waveforms and on-going activity has been proposed as an alternative where the latency, waveform and amplitude of each ERP are estimated with the assumption that the waveform of each ERP is invariant across trial [12, 13]. In addition, applying a subspace-regularized least square method with Gaussian-shaped basis function to recover the ERP waveform from each trial has been suggested in [14].

Single-trial single-channel extraction methods remain important in ERP studies. However, as modern EEG recordings are often flooded with large amounts of data from different electrodes on the scalp, it is increasingly important to utilise the available spatial information to provide a better and more accurate extraction of the ERP-of-interest. For this reason, different multi-channel ERP techniques have been proposed to fully utilize the signals from all available EEG channels. A few notable techniques in this area include: multi-channel estimation using predefined template [15, 16, 17], principal component analysis (PCA), sparse decomposition [18], and independent component analysis (ICA) [19]. Of these techniques, ICA is perhaps one of the more successful techniques and has been used extensively in multi-channel ERP analysis. Its successful use in single-trial ERP analysis has been demonstrated in [20, 21, 22]. Furthermore, the implementation of ICA for ERP extraction also contributed to a remarkable breakthrough in early P300-BCI application [23]. However, the extraction of ERP using ICA is not always straightforward because traditional ICA recovers all of the independent source signals from the mixtures simultaneously. After decomposition, a further source selection is still required before the desired ERP can be studied. Although several refinements towards traditional ICA have been suggested, such as limiting the number of channels [24], performing clustering after ICA decomposition [25, 26] and isolating irrelevant sources based on prior information [27], the main shortcoming of these methods is that full ICA decomposition is still required while the selection of sources are made *a posteriori*.

Since not all source signals are eventually of interest, one of the increasingly popular methods

is the application of constrained ICA (cICA). cICA is an extension of traditional ICA. It attempts to perform signal extraction and source selection jointly by adding prior information as a constraint. As such, it enables ICA to output one or a subset of relevant independent source signals directly. cICA comes in various forms depending on the way a prior information is integrated as a constraint. In [28, 29], a general spatially-constrained ICA (scICA) framework is proposed, based on modifications to traditional ICA. The only difference is that during initialization and re-estimation, there are additional steps where prior spatial information is directly injected to either construct or alter a specific column in a demixing matrix. The scICA is particularly useful when spatial information of a source signal is well-defined. As such, during source estimation, it is not subject to waveform-based spatial information updates on demixing matrix. An example of the use of scICA is removing eye-blink signal from raw EEG signals [28, 29]. However, although scICA can output the desired source signal directly, the method still requires full ICA decomposition which is time-consuming and thus less desirable.

In contrast, other cICA methods that avoid full ICA decomposition have also been proposed. In [30], a method called Functional Source Separation (FSS) was developed to extract single-trial visual evoked potentials from an EEG-fMRI dataset. In FSS, by knowing that the desired ERP occurs at a specific time region, a new constraint function is developed to ensure the extracted source signal has higher averaged amplitude at the predefined time region when compared to the pre-stimulus baseline. However, in FSS, this constraint is added ad hoc and not optimised. In [31], a general temporally-constrained ICA framework is developed to utilize prior temporal information to guide the extraction. The temporal information usually comes in the form of reference signals so that the closeness between an extracted source signal and the reference signals can be computed and further used to constrain the outcome of an ICA. For this reason, temporally-constrained ICA is also known as ICA-with-reference (ICA-R) [31, 32]. ICA-R has been successfully applied in various biomedical signal processing applications, such as artifact removal [32] and rhythmic activities extraction [33, 34].

Since it is usually a requirement that the extracted ERP signal exhibits certain characteristics in its temporal waveform, this paper presents an implementation of one-unit ICA-R that employs a designated reference signal to extract the desired ERP directly. The advantages of using the proposed one-unit ICA-R are: (i) The method is data-driven and only requires time-range of a desired ERP as input,

(ii) It is computationally more attractive since full ICA decomposition is not required.

The organization of this paper is as follows; the procedure for implementing the one-unit ICA-R is described in Section 2.2. As the successful application of ICA-R relies heavily on the design of the reference signal, new ways of generating the reference signal are described in Section 2.3. In Section 3, the performance of the one-unit ICA-R with the proposed reference signals is examined in a simulation study. Section 4 then examines the proposed ICA-R on a real ERP dataset collected from the traditional oddball paradigm before conclusions are drawn in Section 5.

## 2. Proposed method

### 2.1. Traditional ICA

Traditional ICA is a blind source separation technique which aims to recover a set of source signals  $\mathbf{s}(t)$  from a set of observed signals  $\mathbf{x}(t)$  [35]:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) \quad (1)$$

where  $\mathbf{x}(t)$  is a  $N_e \times 1$  vector representing measurements from  $N_e$  channels,  $\mathbf{s}(t)$  is a  $N_s \times 1$  vector representing  $N_s$  EEG source signals and  $\mathbf{A}$  is a mixing matrix with dimension  $N_e \times N_s$ . Here, the source signals  $\mathbf{s}(t)$  is assumed to have zero-mean and with unit variance.

Suppose the EEG source signals are all temporally independent<sup>‡</sup> and spatially-fixed such that the mixing matrix  $\mathbf{A}$  is time-invariant [19, 36]. The objective of ICA is to find a time-invariant  $N_s \times N_e$  demixing matrix  $\mathbf{W}$  such that the recovered source signals  $\mathbf{y}(t)$  are maximally independent:

$$\mathbf{y}(t) = \hat{\mathbf{s}}(t) = \mathbf{W}^\top \mathbf{x}(t) \quad (2)$$

Most traditional ICA performs the extraction by first decomposing the raw EEG signals simultaneously into a number of source signals that equal the number of EEG channels. For EEG recordings that have a large number of EEG channels, this means the extraction and selection of the desired source signals can be a time-consuming and complex process.

### 2.2. One-unit ICA-R

In situations where the ERP-of-interest is known beforehand, it seems a more efficient method is to apply the one-unit ICA-R since it requires only one reference signal and hence yields only one extracted signal [31].

In the literature, the ICA implementation that aims to extract only one source signal is usually known

<sup>‡</sup> Temporal independence means that there is no higher-order as well as second-order correlations between the time courses of two or more different signals.

as one-unit ICA [37]. The optimization problem that defines the one-unit ICA is formulated as follows:

$$\max_{\mathbf{w}} J(\mathbf{w}) = [\langle G(y(t)) \rangle - \langle G(v(t)) \rangle]^2 \quad (3)$$

where  $J(\mathbf{w})$  is a contrast function known as approximated negentropy as proposed by [35],  $y(t) = \mathbf{w}^\top \mathbf{x}(t)$  represents the extracted signal,  $\mathbf{w}$  is the spatial filter,  $v(t)$  is a zero mean and unit variance white Gaussian random signal,  $G(\cdot)$  is a convex non-quadratic function (a number of practical functions for  $G(\cdot)$  are also suggested by [35]) and the time average operator  $\langle z(t) \rangle$  is given by  $\frac{1}{N_t} \sum_t z(t)$  where  $N_t$  is the total number of time samples.

However, when using the above one-unit ICA method, the extracted signal, which is maximally independent, is not guaranteed to be the desired source signal [31]. Thus, to guide the extraction, some prior information of the desired signal is required. One of the prior information that can be provided is a reference signal  $r(t)$ . This subsequently leads to the one-unit ICA-R [31]. The constrained optimization problem that defines the one-unit ICA-R is formulated as follows:

$$\begin{aligned} \max_{\mathbf{w}} \quad & J(\mathbf{w}) = [\langle G(y(t)) \rangle - \langle G(v(t)) \rangle]^2 \\ \text{subject to} \quad & g(\mathbf{w}) = \varepsilon(y(t), r(t)) - \xi \leq 0 \\ & h(\mathbf{w}) = \langle y^2(t) \rangle - 1 = 0 \end{aligned} \quad (4)$$

where  $\varepsilon(y(t), r(t))$  is a closeness measure between  $y(t)$  and  $r(t)$  while  $\xi$  defines the threshold of the closeness measure, and  $h(\mathbf{w})$  is a constraint that ensures the output  $y(t)$  has unit variance. In [31], the constrained optimization problem (4) was solved through an augmented Lagrangian approach, with the inequality constraint  $g(\mathbf{w})$  transformed into equality constraint through the introduction of a slack variable. As such, a Newton-like learning algorithm can be derived to solve problem (4).

In general, one-unit ICA-R relies on the closeness constraint  $g(\mathbf{w})$  to force the algorithm to search for an independent source signal that is nearest to a given reference signal. As the iterative learning process proceeds, the closeness constraint  $g(\mathbf{w})$  is relaxed and eventually becomes inactive whenever the estimate of the independent source signal is drawn into the region where the distance between the estimated source signal and reference signal is less than threshold  $\xi$  [31]. For this reason, besides the design of a reference signal,  $\xi$  is an important parameter that ensures ICA-R converged to the correct source signal. The value of  $\xi$  depends on the choice of the closeness measure  $\varepsilon(y, r)$ . In the case where mean square error (MSE) is used, if  $\xi$  is chosen to be too large, there may exist multiple source signals whose closeness measures fulfill the threshold condition. As a result, the ICA-R may not converge to the desired source signal successfully. In contrast, if  $\xi$

is too small, the algorithm may fail to converge. Thus,  $\xi$  should be selected carefully to ensure the one-unit ICA-R works effectively [31, 33].

In this work, the following measures were taken when one-unit ICA-R is applied:

- Mean square error (MSE) is chosen as the closeness measure where  $\varepsilon(y(t), r(t)) = \langle [y(t) - r(t)]^2 \rangle$ . As detailed in Section 3.2.5, we observed in our case studies that the threshold does not have any significant impact on the performance of one-unit ICA-R. Unless mentioned, the threshold  $\xi$  is selected to be  $10^{-3}$  throughout this paper.
- $G(\cdot) = \log(\cosh(\cdot))$  is selected as the convex non-quadratic function [35].
- Centering and pre-whitening are applied as a signal pre-processing step throughout this study before applying ICA-R [35]. The reason for applying pre-whitening is that it reduces the number of parameters to be estimated and can make the algorithm converged faster [34].
- The scalp distribution, or mixing vector  $\mathbf{a}$  of the extracted signal $\S$  is estimated through the minimum mean square error as [38, 39]:

$$\hat{\mathbf{a}} = \frac{\langle \mathbf{x}(t)y(t) \rangle}{\langle y^2(t) \rangle} \quad (5)$$

### 2.3. New reference signal design for ERP applications

The reference signal plays a vital role in one-unit ICA-R. It has to be constructed carefully so that it incorporates the appropriate prior information. Previous studies have shown that the reference signal is not required to be exactly similar to the desired signal, provided that it carries the important characteristics of the desired signal. One of the common methods to generate the reference signal is to create a synthetic signal that represents the time location of the desired signal. For example, to extract rhythmic activities, such as fetal ECG from the maternal ECG, a rectangular pulse train whose frequency is similar to the fetal ECG can be used [34]. Alternatively, if the desired signal is strong and readily observed in a certain EEG channel, a reference signal can be constructed by converting the signal segment that exceeds a given detection threshold into a series of rectangular pulses that represent the time location of the desired signal [32]. Unfortunately, these methods are not applicable for ERPs, since ERPs are weak and

$\S$  In traditional ICA, under the square mixing assumption (i.e.  $N_e = N_s$ ), the resultant  $\mathbf{W}$  is an invertible square matrix that also represents the mixing matrix  $\mathbf{A}$  when it is inverted. The column vectors of  $\mathbf{A}$  potentially convey the topographic information of a source signal. However, unlike traditional ICA, this scalp distribution of the extracted source signal cannot be obtained by inverting  $\mathbf{w}$  from the ICA-R.

non-periodic. Therefore, a new way of generating the reference signal is required for ERP application.

**2.3.1. Fixed reference signal** The major problem in generating a reference signal for the desired ERP is to determine the ERP time region of each individual trial. In practice, these ERP time regions are very difficult to obtain even by a simple threshold. Thus, one of the more feasible solutions is to estimate the ERP time region from the trial-averaged signal and assume that the desired ERP in all trials reside mainly in the estimated time region.

The first reference signal  $r_1(t)$  proposed for our one-unit ICA-R is constructed based on the above idea. Suppose the input signal  $\mathbf{x}(t)$  of the one-unit ICA-R is taken from a large segment of the multi-channel EEG recording and may contain target and non-target trials. Let  $t_k$  be the stimulus onset of the  $k$ -th target trial, and the time interval  $[t_{start}, t_{end}]$  is the ERP time region estimated from the trial-averaged target signal. The first reference signal  $r_1(t)$  is constructed by generating a series of rectangular pulses on  $[t_{start}, t_{end}]$  after the stimulus onset  $t_k$  of each trial:

$$r_1(t) = \begin{cases} 1 & \text{for } t \in \mathcal{T} \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

where the set  $\mathcal{T} = \bigcup_k [t_k + t_{start}, t_k + t_{end}]$  represents the overall estimated ERP time region [See Fig. 1(a)].

**2.3.2. Discriminative-based reference signal** The underlying ERPs are unlikely to be stationary across trials. As a result, the mismatch information in  $r_1(t)$  may prevent one-unit ICA-R from extracting the desired source signal accurately. Since the ERPs are non-stationary, we proposed a new type of reference signal to identify the trial-to-trial variations in the ERP.

In EEG recordings, each ERP usually exhibits a unique scalp distribution that occurs at a specific time interval after stimulus onset [3, 4]. Motivated by this consistent finding, we propose the application of a discriminative-based spatial filter to generate the reference signal. In general, the idea of a discriminative-based spatial filter is similar to the work by Blankertz *et al.* [40, 41] and Lee *et al.* [42]. Firstly, a spatial filter is trained to learn the differences between scalp distribution of the desired ERP and the background EEG/ERP. Secondly, by applying the spatial filter on the EEG recordings, a signal that reflects the activation time of the desired ERP is produced. In this work, this signal is used as a reference signal to guide the one-unit ICA-R.

The procedure to obtain the discriminative-based reference signal  $r_2(t)$  is as follows (See Fig. 1(b)):

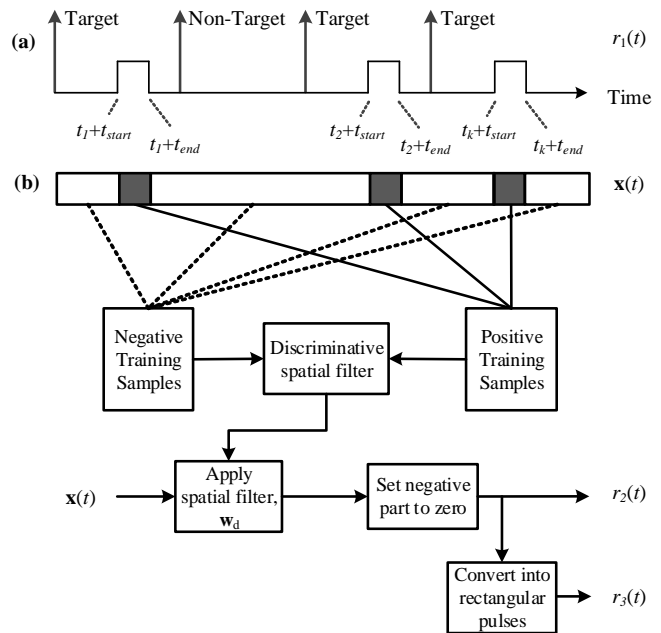


Figure 1: Procedure for generating different types of reference signal for ERP extraction. (a) the fixed reference signal  $r_1(t)$ , (b) the discriminative-based reference signal  $r_2(t)$  and its reshaped version  $r_3(t)$ . Note that the negative training samples do not necessarily have to be taken only from non-target condition.

- (i) Those multi-channel measurement vectors with time indices within the set  $\mathcal{T}$ , i.e.  $\mathbf{x}(t)|_{t \in \mathcal{T}}$ , are taken as the positive training samples while the remaining measurement vectors are taken as the negative training samples.
- (ii) Let the sets  $\mathcal{T}_1$  and  $\mathcal{T}_2$  contain, respectively, the time indices of the positive class and those of the negative class where  $\mathcal{T}_1 = \{t \in \mathcal{T}\}$  and  $\mathcal{T}_2 = \{t \notin \mathcal{T}\}$ . Let  $T_1$  and  $T_2$  be the number of elements in  $\mathcal{T}_1$  and  $\mathcal{T}_2$ , respectively. The two  $N_e \times 1$  vectors which represent the mean spatial distribution of the positive and negative classes can be computed as:

$$\boldsymbol{\mu}_c = \frac{1}{T_c} \sum_{t \in \mathcal{T}_c} \mathbf{x}(t), \quad c = 1, 2 \quad (7)$$

and the  $N_e \times N_e$  covariance matrix for both classes as:

$$\mathbf{S}_c = \frac{1}{T_c} \sum_{t \in \mathcal{T}_c} (\mathbf{x}(t) - \boldsymbol{\mu}_c)(\mathbf{x}(t) - \boldsymbol{\mu}_c)^\top, \quad c = 1, 2 \quad (8)$$

- (iii) Based on these training samples, the discriminative spatial filter  $\mathbf{w}_d$  can be obtained as:

$$\mathbf{w}_d = \mathbf{S}_w^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) \quad (9)$$

where by accounting for the prior class probability, the within-class scatter matrix  $\mathbf{S}_w$  is calculated as:

$$\mathbf{S}_w = \frac{1}{T_1 + T_2} (T_1 \mathbf{S}_1 + T_2 \mathbf{S}_2) \quad (10)$$

- (iv) By applying the discriminative spatial filter, the discriminative-based reference signal  $r_2(t)$  is given as:

$$r_2(t) = \max(\mathbf{w}_d^\top \mathbf{x}(t) - b, 0) \quad (11)$$

where the threshold of the spatial filter  $b$  is given as:

$$b = \mathbf{w}_d^\top (\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2) / 2 \quad (12)$$

It should be noted that as shown in Fig. 1, the way of training the discriminative-based spatial filter is slightly different from the approach mentioned in [40, 41]. The method in [40, 41] is restrictive in the sense that the negative training samples are taken from the same time region of the opposite experiment condition. This type of discriminative spatial filter can only be trained on certain type of ERPs (e.g. P300 [5]) whereby the desired ERP appears on one condition but not the other. In contrast, the negative training samples in our method are taken from any time region that is not included by  $\mathcal{T}$  regardless of the experiment condition. Thus, the method described in Section 2.3.2 can be used on different types of ERPs. In fact, loosely speaking, as shown in Section 4, it can also work on other type of EEG activities such as eye-blinks as long as the desired signal has an unique spatial pattern over certain time region.

### 2.3.3. Binary form of discriminative-based reference signal

In practice, the rectangular pulse train is often selected for constructing the reference signal for ICA-R. Thus, for comparison, a third reference signal  $r_3(t)$  is obtained by reshaping  $r_2(t)$  into rectangular pulses as follows:

$$r_3(t) = \begin{cases} 1 & \text{for } r_2(t) > 0 \\ 0 & \text{otherwise} \end{cases} \quad (13)$$

## 3. Simulation Study

### 3.1. Semi-simulated EEG data

In this section, our first objective is to examine whether the one-unit ICA-R could be used to extract the desired ERP directly. After which, the next objective is to understand which of the three reference signals proposed in Section 2.3 would be most suitable for the one-unit ICA-R.

In practice, examination of the ERPs is usually challenging because 1) ERPs vary across trials, 2) they are easily contaminated by background EEG activities and 3) the recorded dataset is usually a

massive spatio-temporal matrix that contains at least 32 EEG channels and span over at least 50 trials. To replicate this scenario, the simulation study will be carried out on semi-simulated ERP datasets which are generated by mixing synthetic ERPs with real EEG background noises. Mathematically, the linear EEG model for generating the semi-simulated ERPs is given as follows:

$$\begin{aligned} \mathbf{x}(t) &= \alpha \mathbf{x}_{\text{ERP}}(t) + \mathbf{x}_{\text{EEG}}(t) \\ &= \alpha \left\{ \sum_{i=1}^4 \mathbf{a}_i^\top s_i(t) \right\} + \mathbf{x}_{\text{EEG}}(t) \end{aligned} \quad (14)$$

where  $\mathbf{a}_i$  is the mixing vector,  $s_i(t)$  is the simulated source signal of  $i$ -th ERP and  $\alpha$  is a scaling factor that adjusts the channel-averaged signal-to-noise ratio (SNR) between  $\mathbf{x}_{\text{ERP}}(t)$  and  $\mathbf{x}_{\text{EEG}}(t)$ . The background EEG noise  $\mathbf{x}_{\text{EEG}}(t)$  is taken from a recording of a subject who was instructed to sit and watch a movie for 40 mins.

ERPs often appear as a series of positive and negative peaks, with each peak representing different scalp distribution and appearing at different time range following stimulus onset. For this reason, firstly, assuming a sampling frequency of 1 kHz, four source signals  $s_i(t)$  are generated to represent four different types of ERPs. Each source signal contains a series of 100 ERPs, with each ERP appearing in every 1000 time samples. In other words, the inter-stimulus interval is assumed to be 1000 ms. Secondly, every ERP in each  $s_i(t)$  is generated as a Gaussian pulse:

$$s_{i,k}(t) = c_{i,k} \exp \left\{ -\frac{(t - \tau_{i,k})^2}{2\delta_{i,k}^2} \right\} \quad (15)$$

where  $c_{i,k}$ ,  $\tau_{i,k}$  and  $\delta_{i,k}$  are the peak amplitude, latency and width of the  $k$ -th ERP respectively in  $i$ -th source signal. Table 1 summarizes the parameters used to generate different ERPs, where  $\sigma_\tau$  is the latency standard deviation of ERPs in  $s_4(t)$ , while  $[v_1, v_2]$  denotes a uniformly distributed variable between value  $v_1$  and  $v_2$ .

Table 1: Parameters used to generate different ERP source signals  $s_i(t)$

Parameters	$s_1(t)$	$s_2(t)$	$s_3(t)$	$s_4(t)$
$c_{i,k}$	[0, 1]	[-1, 0]	[0, 1]	[0, 1]
$\tau_{i,k}$ (ms)	$50 \pm 5$	$150 \pm 5$	$250 \pm 5$	$650 \pm \sigma_\tau$
$\delta_{i,k}$ (ms)	[10, 20]	[10, 20]	[10, 20]	[50, 100]

In terms of  $\mathbf{a}_i$ , the mixing vector for every ERP is fixed across trials. However, in each simulated dataset, they are randomly generated with uniform distribution [0,1] to represent different scalp distributions. The parameters in Table 1 are selected in such a way that

$s_1(t)$ ,  $s_2(t)$  and  $s_3(t)$  represents the sensory ERPs which tend to appear earlier after stimulus onset, have narrower waveform and smaller latency variations while  $s_4(t)$  represents the cognitive ERP which tend to appear much later and has broader waveform [1].

In terms of  $\mathbf{x}_{\text{EEG}}(t)$ , it was collected from a 40-channel NuAmps EEG system (34 EEG channels, 2 reference channels and 4 EOG channels). As a pre-processing step, the EEG signals were bandpass filtered between 0.5 Hz - 35 Hz while eye-blinks were removed using the method of [32] by taking the electrode VEOG as a reference. For each simulation run, a semi-simulated EEG segment  $\mathbf{x}(t)$  was generated by adding  $\mathbf{x}_{\text{ERP}}(t)$  into the real EEG signal  $\mathbf{x}_{\text{EEG}}(t)$ , with channel-averaged SNR between  $\mathbf{x}_{\text{ERP}}(t)$  and  $\mathbf{x}_{\text{EEG}}(t)$  adjusted based on the type of the experiment. For evaluation, the performance of the proposed ICA-R is examined based on its ability to extract the ERP source signal  $s_4(t)$ . The main reason of choosing source signal  $s_4(t)$  is that the ERPs in  $s_4(t)$  are further away from those ERPs in  $s_3(t)$  so it allows us more flexibility in adjusting  $\sigma_\tau$  and further study the impact of latency variations have on the design of the reference signals.

Two main experiments were carried out in this study. Experiment 1 involves testing 11 different values of  $\sigma_\tau$  ranging from 0 ms to 100 ms, with SNR set to 0 dB. Meanwhile, Experiment 2 involves testing 7 different levels of SNR, ranging from -15 dB to 15 dB, with  $\sigma_\tau$  set to 50 ms. In these two experiments, 250 different semi-simulated signals were generated for each value of  $\sigma_\tau$  and SNR.

To recover the simulated source signal, the one-unit ICA-R was applied to each semi-simulated signal with the three different reference signals  $r_1(t)$ ,  $r_2(t)$  and  $r_3(t)$ . During the training of the one-unit ICA, the estimated ERP time region  $\mathcal{T}$  was set between 575 ms and 725 ms after each stimulus onset. The extraction quality of the proposed one-unit ICA-R was examined by measuring the performance index (PI) as defined below [33]:

$$\text{PI} = -10\log_{10}\langle [s_4(t) - y(t)]^2 \rangle \quad (16)$$

To use this metric, both the simulated source signal  $s_4(t)$  and extracted signal  $y(t)$  were normalised to be zero mean with unit variance. This metric can be seen as the inverse of the MSE that describes the similarity between the simulated source signal and the extracted signal in decibel (dB). Generally, values larger than 20 dB indicate good extraction quality [33]. In this study, to serve as a baseline comparison, three different traditional ICA methods namely extended-Infomax ICA [45, 44], negentropy-based FastICA [35] and JADE [46] were selected. The implementation of Infomax and JADE ICA methods are taken from the well-known EEGLAB toolbox [47] while FastICA method was taken from FastICA toolbox [48]. For

traditional ICA, their performances were evaluated by choosing the source signal with the highest PI||.

### 3.2. Results and discussion

**3.2.1. Qualitative assessment** Fig. 2 illustrates the example of signals at different stages of the simulation study. Fig. 2(a)–(d) represents the simulated source signals  $s_1(t)$ ,  $s_2(t)$ ,  $s_3(t)$  and  $s_4(t)$ . Fig. 2(e) shows the semi-simulated mixed signals  $\mathbf{x}(t)$ . Fig. 2(f)–(h) show the reference signals  $r_1(t)$ ,  $r_2(t)$  and  $r_3(t)$  generated for this set of mixed signals, while their corresponding extracted signals  $y(t)$  with ICA-R are shown in Fig. 2(j)–(l) respectively. In addition, Fig. 2(i) represent the targeted source signal  $s_4(t)$  which is normalised to have zero-mean and unit variance. From these figures, it can be observed that one-unit ICA-R is robust against the design of the reference signal, since the results show that all reference signals managed to guide the one-unit ICA-R to extract the desired source signal effectively. However, in terms of the amplitude range, the extracted signals vary slightly from the original source signals, mainly because of the scaling ambiguity and the zero-mean and unit-variance source signal assumption [35]. In any event, the extracted signals closely resemble the normalised source signal, as observed in Fig. 2(j)–(l). In addition, from these figures, it can be observed that the signal from  $r_2(t)$  resembles the normalised source signal the closest. By examining the temporal waveform of each reference signal, it can be seen that the performance difference in extraction is caused mainly by the characteristic of the reference signal, as shown in Fig. 2(f)–(h).

To examine the characteristic of each reference signal in greater detail, 2D graphs were plotted in Fig. 3 to reveal how each reference signal acts at the single-trial level. By dividing the signals into trials and sorted them according to the peak latencies, each plot in Fig. 3(a) shows one of the 250 simulated source signals  $s_4(t)$  with varying peak amplitudes, latencies and widths. The corresponding responses given by each reference signal were shown in Fig 3(b)–(d). In Fig. 3(b), it is easily observed that the fixed reference signal  $r_1(t)$  does not change with respect to any amplitude and latency fluctuations. In contrast, due to the ability of the discriminative spatial filter to learn the difference between the relevant ERP and irrelevant EEG/ERP time regions, Fig. 3(c) and (d) show that both reference signals  $r_2(t)$  and  $r_3(t)$  managed to pick up the changes in the ERP peak latencies and adapted their waveforms accordingly.

|| Traditional ICA decomposes the ERP data into a number of source signals equal to the number of EEG channels. As a result, source selection is necessary before the ERP-of-interest can be studied.

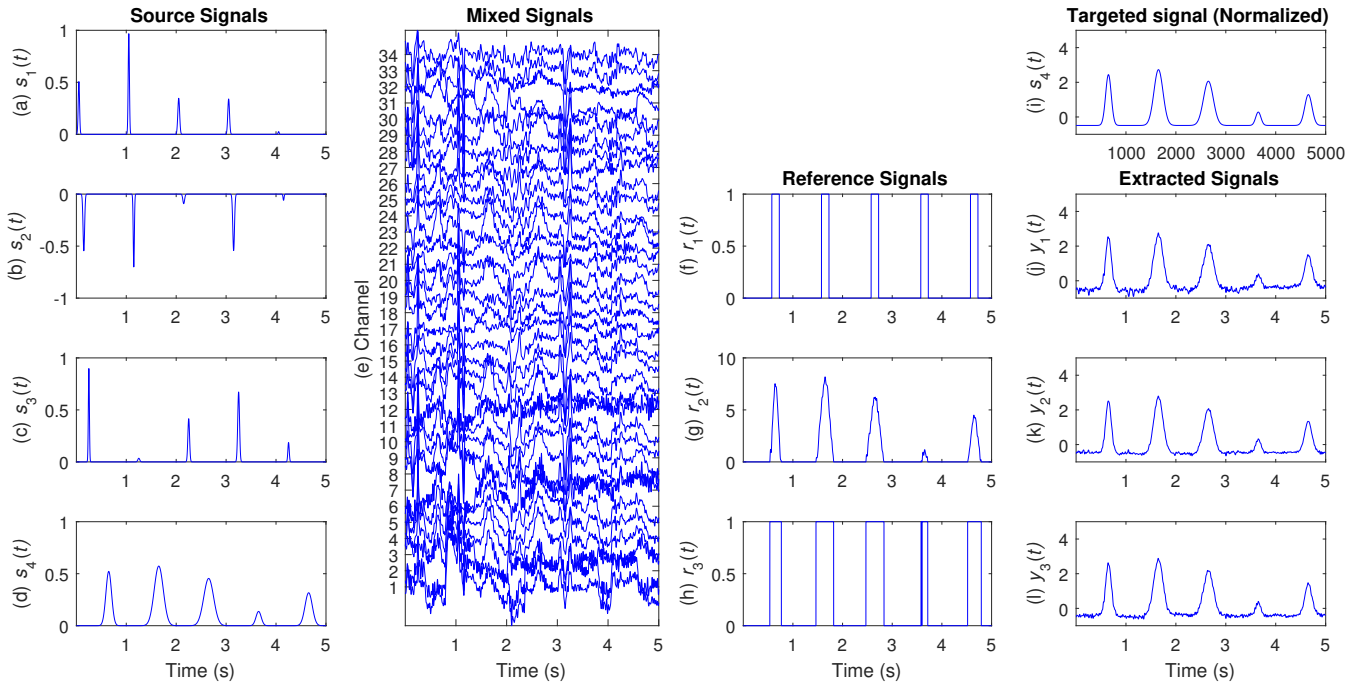


Figure 2: (a)–(d) a segment of source signals  $s_1(t)$ ,  $s_2(t)$ ,  $s_3(t)$ ,  $s_4(t)$ ; (e) semi-simulated signal  $\mathbf{x}(t)$ ; (f)–(h) the reference signals that were used to recover  $s_4(t)$ ; (i) the normalized targeted signal  $s_4(t)$ ; (j)–(l) the resultant signals  $y(t)$  extracted from ICA-R

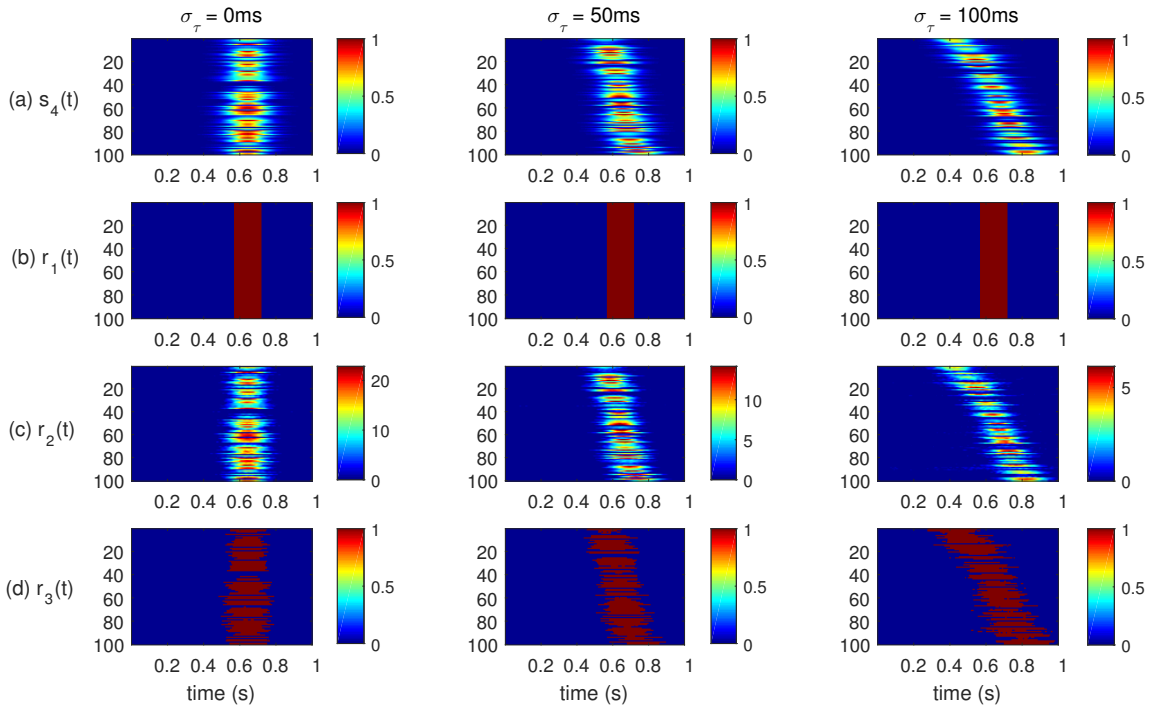


Figure 3: Visualization of different signals sorted at single-trial level. (a) The examples of simulated ERPs with varying peak amplitude, width and also different levels of latency variation  $\sigma_\tau$  (plotted in different columns); (b)–(d) The responses given by the reference signal  $r_1(t)$ ,  $r_2(t)$  and  $r_3(t)$  respectively. For all plot, the  $y$ -axis represents the trial number and the  $x$ -axis gives the time in second. The color intensity shows the amplitude of a signal.



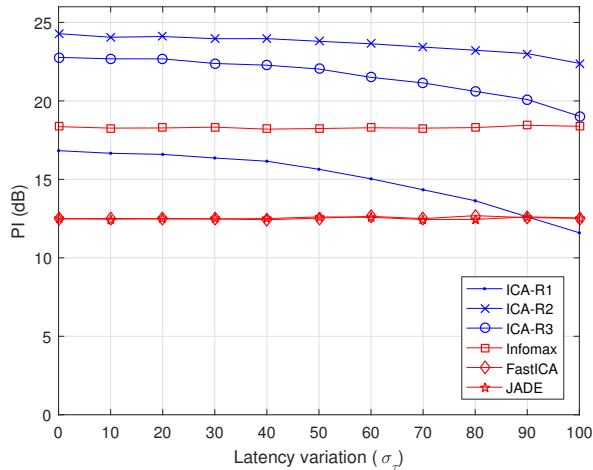


Figure 4: Performance Index  $PI$  of different ICA algorithms under different  $\sigma_\tau$  when SNR = 0 dB

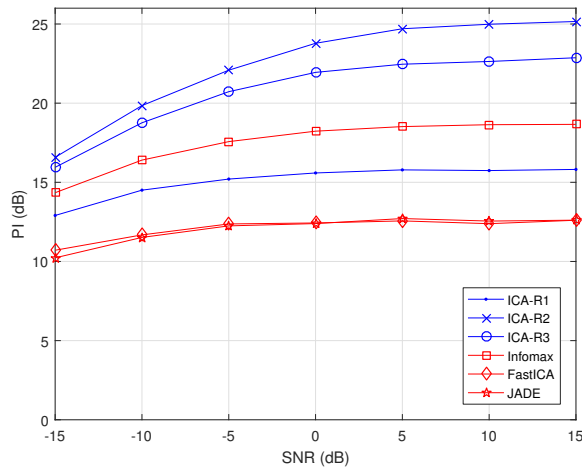


Figure 5: Performance Index  $PI$  of different ICA algorithms under different SNR when  $\sigma_\tau = 50$  ms

**3.2.2. Quantitative examination** Fig. 4 and 5 shows the averaged PI scores achieved by the different ICA methods when different latency variations  $\sigma_\tau$  and SNR are encountered in the ERPs respectively. Meanwhile, the overall PI scores achieved by different methods in these two experiments are tabulated in Table 2 and 3.

When comparing the different reference signals, the results in both Fig. 4 and 5 show that the adaptive ability of reference signals  $r_2(t)$  and  $r_3(t)$  provides the one-unit ICA-R with a better overall performance in extraction. Specifically, in both experiments,  $r_2(t)$  achieved the best performance, followed by  $r_3(t)$  and  $r_1(t)$ . The signal quality extracted by using  $r_2(t)$  is also better when compared to the traditional Infomax-based ICA. The finding that  $r_2(t)$  is superior is similar to the findings based on pure simulated signals [49], in which, the one-unit ICA-R may outperform the traditional ICA when the reference signal is designed

correctly. Interestingly, a comparison between  $r_2(t)$  and  $r_3(t)$  also shows that the conventional approach of using rectangular pulses in the reference signal  $r_3(t)$  does not guarantee a better extraction performance. From the analysis, it can be concluded that it may be important to include the amplitude variation of the ERPs in the design of the reference signal.

Table 2: Average PI under different  $\sigma_\tau$  when SNR = 0 dB

Method	Reference signal	PI (dB)
one-unit ICA-R	$r_1(t)$	$15.04 \pm 1.78$
	$r_2(t)$	$23.63 \pm 0.56$
	$r_3(t)$	$21.56 \pm 1.23$
Infomax ICA		$18.31 \pm 0.07$
FastICA		$12.53 \pm 0.08$
JADE		$12.52 \pm 0.06$

Table 3: Average PI under different SNR levels when  $\sigma_\tau = 50$ ms

Method	Reference signal	PI (dB)
one-unit ICA-R	$r_1(t)$	$15.08 \pm 1.07$
	$r_2(t)$	$22.45 \pm 3.21$
	$r_3(t)$	$20.76 \pm 2.56$
Infomax ICA		$17.48 \pm 1.60$
FastICA		$12.11 \pm 0.69$
JADE		$12.04 \pm 0.89$

To understand the computational efficiency of different ICA methods, a further test is performed by measuring the convergence time spent by each ICA method on the semi-simulated signals under the condition of  $\sigma_\tau = 50$  ms and SNR = 0 dB for 250 times. For this test, a CPU with an Intel Core i7 4770 3.4 GHz processor and 16 Gb memory RAM was used as the test bench. The convergence time required by different ICA methods was tabulated in Table 4.

Table 4: Average convergence time per run when SNR = 0 dB and  $\sigma_\tau = 50$ ms

Method	Reference signal	Seconds
one-unit ICA-R	$r_1(t)$	$0.73 \pm 0.06$
	$r_2(t)$	$0.12 \pm 0.06$
	$r_3(t)$	$0.58 \pm 0.02$
Infomax ICA		$24.04 \pm 51.21$
FastICA		$126.05 \pm 59.81$
JADE		$24.05 \pm 5.07$

As shown in Table 4, for simulated data that contain 34 channels and 100,000 time samples, all three variants of the proposed one-unit ICA-R only takes an

average of less than 0.8 s to converge and complete the extraction. Among the reference signals, the ICA-R that uses  $r_2(t)$  was the most efficient, followed by  $r_3(t)$  and  $r_1(t)$ . In contrast, all traditional ICA methods took longer time on the average to converge and most importantly, these times do not include actual time required for source selection in practice. It should also be noted that there are large differences in the convergence time between different traditional ICA methods. Apart from the computation complexity, these differences could be partly due to the way a ICA method is implemented in MATLAB. As reported in [47], the Infomax ICA method implemented in EEGLAB is actually optimised in terms of coding and also the way the solution converged.

In summary, from the performance and convergence time test, one can conclude that the use of one-unit ICA-R for ERP extraction is generally possible with a proper reference signal design. A reference signal that resembles more the desired signal, such as  $r_2(t)$  will help ICA-R to converge quicker and also better extract source signals. In addition, compared to the traditional ICA that performs full source extraction, one-unit ICA-R requires far less time, since the extraction involves only one source signal. Most importantly, the extraction of the desired ERP is straightforward without human intervention.

**3.2.3. Effect of the specified time region  $\mathcal{T}$**  In Fig. 4, the results show that compared to the traditional ICA, the performance of all the one-unit ICA-R decreases when the latency variation of the ERPs increases. This indicates that apart from the choice of reference signals, another factor that affects the performance of the one unit ICA-R is the overall estimated time region  $\mathcal{T}$ . In practice, there will be situations when the time region of a specific trial is not covered well by  $\mathcal{T}$ . When this situation occurs, it introduces mismatch between  $r_1(t)$  and the original source signal, and subsequently reduces the performance of the corresponding one-unit ICA-R. Likewise, for  $r_2(t)$  and  $r_3(t)$ , when  $\mathcal{T}$  introduces the mislabeled training samples (i.e. the positive training samples are not taken from the correct ERP time region of each respective trial and vice versa), the performance of the one-unit ICA-R is also reduced. However, from Fig. 4, a closer examination reveals that among all three reference signals, the performance of  $r_2(t)$  is less affected by the mismatch, with only  $\sim 2$  dB difference between PI at  $\sigma_\tau=0$  ms and  $\sigma_\tau=100$  ms. This is mainly because as can be seen by comparing Fig. 3(a) and (c), the reference signal  $r_2(t)$  is robust and is able to adapt to the amplitude, width and latency variations of the underlying ERPs.

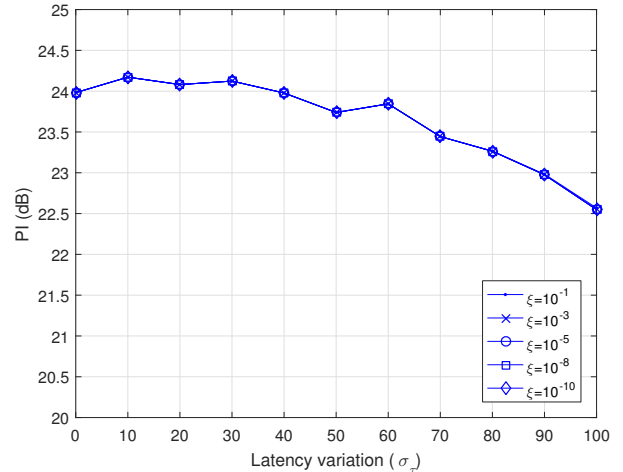


Figure 6: Performance Index  $PI$  of one-unit ICA-R with  $r_2(t)$  when different threshold values were used

**3.2.4. Effect of the SNR between the ERP and background EEG** Fig. 5 show the performance of different ICA methods are affected by the level of background EEG activities. Overall, all the ICA methods have better performance when the SNR is higher, but deteriorates rapidly when the SNR is smaller than 0 dB. In this test, it is observed that the proposed ICA-R with  $r_2(t)$  remains the best among all the different ICA methods. Specifically, it is the only method that achieves 20 dB performance when the SNR is at -10 dB. The reason why the proposed ICA-R has a good performance under the noisy condition is that the discriminative-based spatial filter is a statistical method. As shown in Eq. 9, in order to obtain a robust estimate of the scalp distribution of the desired ERP, it minimises the random noises by performing a time-averaging of scalp distribution  $\mathbf{x}(t)$  across the relevant EEG/ERP segment. As a result, the method is able to generate a good reference signal for the one-unit ICA-R since the estimated scalp distribution  $\mathbf{u}_1$  still resembles the actual scalp-distribution of the desired ERP.

**3.2.5. Effect of different thresholds on the ICA-R** In this section, Experiment 1 is repeated by using different threshold values  $\xi$  for one-unit ICA-R with  $r_2(t)$ . Fig. 6 shows that one-unit ICA-R achieved similar PI scores under different values of  $\xi$ . This result implies that the proposed ICA-R does not face any difficulties in converging to the correct source signal at different threshold levels. This is because there is no other signal except the target source signal that is similar to the reference signal. In practice, the

¶  $\mathbf{x}(t)$  represents measurements from multiple channel at time index  $t$ . As a result, it also represents a scalp distribution at time index  $t$ .

performance of ICA-R is also likely to be insensitive to the choices of the threshold level since every ERP has a specific activation time and thus a unique source signal that is rarely similar to other ERPs.

## 4. Real ERP Study

### 4.1. Dataset

The objective in this section is to examine whether the one-unit ICA-R can extract the desired ERP in the real ERP dataset. It should be noted that since the ground truth of the real ERP source signal is not available, it is not possible to measure the quantitative performance of different ICA methods. Thus, the performance of the proposed ICA-R method will be examined qualitatively based on the following three questions: 1.) does the proposed ICA-R method managed to extract the desired ERP correctly?, 2.) does the extracted signal similar to the results from traditional ICA methods especially Infomax ICA since it has the better performance in the simulation study? and 3.) is the proposed ICA-R able to extract the ERP from trials that are heavily-contaminated by artifacts such as eye-blinks?

In this section, the application of the proposed one-unit ICA-R was evaluated on a publicly available P300 ERP dataset contributed by Ting *et al.* [50, 51]. This dataset consists of ten participants and was originally collected for a study that utilises P300 ERP to provide a quantitative assessment between 2D and 3D visualization modalities. The procedure used by Ting *et al.* [51] for collecting this dataset is as follows. Each participant was randomly presented with an image of a cube and sphere for 100 ms, followed by a black screen for 1,000 ms. The number of cubes and spheres used in this experiment were 320 and 80 respectively, with the cube treated as non-target and the sphere treated as target. Whenever a sphere was presented, the participant was required to respond by pressing a key on a keyboard. For this dataset, 34 channels of EEG were recorded using Neuroscan Nuamps system with a sampling frequency of 1,000 Hz.

### 4.2. Procedure

For each EEG recording, the EEG segment from 0.2 s before the first stimulus to 1 s after the last stimulus was extracted for the proposed one-unit ICA-R. This segment of EEG signals contains all the target and non-target trials. The goal of using the one-unit ICA-R is to extract a source signal that reflects only P300 from this EEG segment such that after dividing it into epochs, the source signal should only contain P300 in the target trial and no activity for the non-target trials.

The data preprocessing step was as follows. First, the EEG segment was referenced to the average of channel A1 and A2 before being bandpass-filtered<sup>+</sup> between 0.5 - 30 Hz. Any problematic EEG channels were removed from the analysis.

In terms of  $\mathcal{T}$ , there are various ways that can be used to estimate the start  $t_{start}$  and end time  $t_{end}$  of the P300. As shown in Fig. 2, as long as the selected time region covers most of the P300 at trial-averaged signal, the one-unit ICA-R can perform effectively even if the selected time region may not necessarily reflect the exact time region of P300 at single-trial level. To estimate  $t_{start}$  and end time  $t_{end}$ , a direct approach is by examining the spatio-temporal patterns of the trial-averaged target signal and visually determine the start and end time of the P300 scalp distribution.

Alternatively, the start  $t_{start}$  and end time  $t_{end}$  can also be estimated by applying segmentation techniques on the spatio-temporal patterns of the trial-averaged target signal [52, 53]. In this study, a 2-cluster K-means algorithm is used [54, 55]. Specifically, by taking advantages that P300 ERP only occurs on one experiment condition but not the other, a multi-channel differential signal is computed by subtracting the trial-averaged non-target signal from the trial-averaged target signal. The segmentation algorithm is then performed on the differential signal to divide the signal into two time regions (i.e. “active” and “quiet” regions). The start  $t_{start}$  and end time  $t_{end}$  of P300 is selected based on the start and end of the “active” region.

In each recording, the start  $t_{start}$  and end time  $t_{end}$  of P300 was first estimated by the segmentation algorithm. Those measurement vectors  $\mathbf{x}(t)$  that are within the range of  $[t_{start}, t_{end}]$  after target trials onset were taken as positive training samples while the remaining  $\mathbf{x}(t)$  were taken as negative training samples for generating a discriminative-based reference signal as discussed in Section 2.3.2. After that, the reference signal is used together with the proposed one-unit ICA-R to extract a source signal. For analysis, the resultant extracted signal was divided into single-trial target and non-target epochs, with each epoch containing a 0–1,000 ms EEG segment after stimulus onset. Here, the target group represents the correctly identified target trials. Baseline correction was performed using a 200 ms time segment prior to the stimulus onset. For comparison, the traditional ICA method was applied to the given EEG segment whereby the source signal relevant to the P300 ERP was manually selected for comparison.

<sup>+</sup> A combination of a forward-backward 2<sup>nd</sup> order elliptic highpass and 8<sup>th</sup> order elliptic lowpass filters.

### 4.3. Results

#### 4.3.1. Case Study 1: Minimally-contaminated trials

In this dataset, since the subjects were not restricted from blinking their eyes, the EEG recordings from different subjects suffer varying degree of eye-blinks contamination. In this section, an EEG recording that has less than 10% of the trials affected by eye-blinks was selected for testing the proposed ICA-R.

Fig. 7 shows the results of applying different ICA algorithms on the selected EEG recording where the first, second, third and fourth columns reflect respectively the trial-averaged target/non-target signals, scalp distribution, single-trial target and non-target epochs extracted from different ICA methods. The target trials in third column are sorted based on the peak latencies of P300 ERP detected from the ICA-R with  $r_2(t)$  method.

In the first row, by examining raw EEG signals from Channel Pz, the result shows the presence of a P300 ERP around 425 ms in the trial-averaged target signal (blue line). The activation of other early ERPs can also be observed around 200 ms in both trial-averaged target and non-target signals. When examining the scalp distribution of P300 taken from trial-averaged target signal at  $t=425$  ms, the result reveals that P300 originates from the centro-parietal region, which is consistent with the literature [5]. The second, third and fourth rows of Fig. 7 show the results of applying the proposed one-unit ICA-R with  $r_1(t)$ ,  $r_2(t)$  and  $r_3(t)$  respectively. The results from different one-unit ICA-R methods are similar in terms of waveform and scalp distribution. By examining the trial-averaged target signals, all one-unit ICA-R methods managed to extract the P300 ERP and suppress the activity of other ERPs effectively. The signals from all ICA-R methods have the centro-parietal distribution\* which is similar to the one obtained from raw EEG signals. The results at single-trial level also reveals that all one-unit ICA-R methods managed to capture the P300 ERPs consistently between 300 and 700 ms. For comparison, the fifth, sixth and seventh rows in Fig. 7 represents the source signals selected from Infomax ICA, FastICA and JADE algorithms respectively.

#### 4.3.2. Case Study 2: Heavily-contaminated trials

In this section, an ERP recording that is heavily contaminated by eye-blinks was selected for testing. The first row of Fig. 8 shows the impact of the eye-blinks has on the EEG recording. It can be observed that for 70% of the target and non-target

\* It should be noted that the scalp distribution of an extracted source signal is represented by the mixing vector  $\mathbf{a}$ . For ICA-R,  $\mathbf{a}$  is estimated by using Eq. 5 while for traditional ICA,  $\mathbf{a}$  is taken from one of the columns from mixing matrix.

trials, the subject blinked his/her eyes 250 ms after the trial onset. As a result, instead of randomly-occurred events, the eye-blinks have systematically contaminated the scalp distribution of P300 ERP at 400 ms. Under this condition, it is not possible to perform the proposed ICA-R directly because the discriminative-based spatial filter cannot learn the P300 scalp distribution correctly without removing the eye-blinks.

The traditional method to examine the desired ERP under this condition is to reject the contaminated trials. However, this method is less desired since it reduces the number of usable trials. Another approach is to “bypass”‡ the eye-blinks artifact by performing a full ICA decomposition and select the source signal that represents the desired ERP directly. As mentioned previously, this method is not efficient since it requires higher computational resources and yet not all the extracted source signals are utilized. In addition, the source selection stage often involves human intervention and thus requires domain expertise.

Alternatively, another way of extracting the desired ERP is to perform a sequential extraction using the proposed one-unit ICA-R. Theoretically, one-unit ICA can be extended to perform either a partial or full ICA decomposition through a series of source extraction and deflation [38, 39]. In this section, by exploiting this characteristic of one-unit ICA, two stages of extraction were performed based on the proposed one-unit ICA-R, with the first stage aims at removing the eye-blink artifacts while the second stage aims at extracting the desired ERP from the artifact-free trials.

Ideally, in the first stage, if a dataset contains the signal from channel VEOG, this signal can be used directly as a reference signal to guide the ICA-R to remove the eye-blinks. However, since the selected dataset does not contain channel VEOG, the signals from channel Fp1 and Fp2 were selected since these channels are nearest to the eyes. The method of removing eye-blinks is similar to the method described in [32], except that the discriminative-based reference signal is used.

The steps of generating a discriminative-based spatial filter to remove the eye-blinks are described as follows. Firstly, the signals from channel Fp1 and Fp2 were averaged. Secondly, by applying a threshold of  $50 \mu\text{V}$ , those time indices when the averaged signal exceeds the threshold were selected as the positive training time region while the remaining time indices were selected as the negative training time region.

‡ In full ICA decomposition, the activity of eye-blinks is one of the extracted source signals. However, it is ignored during the source selection stage.

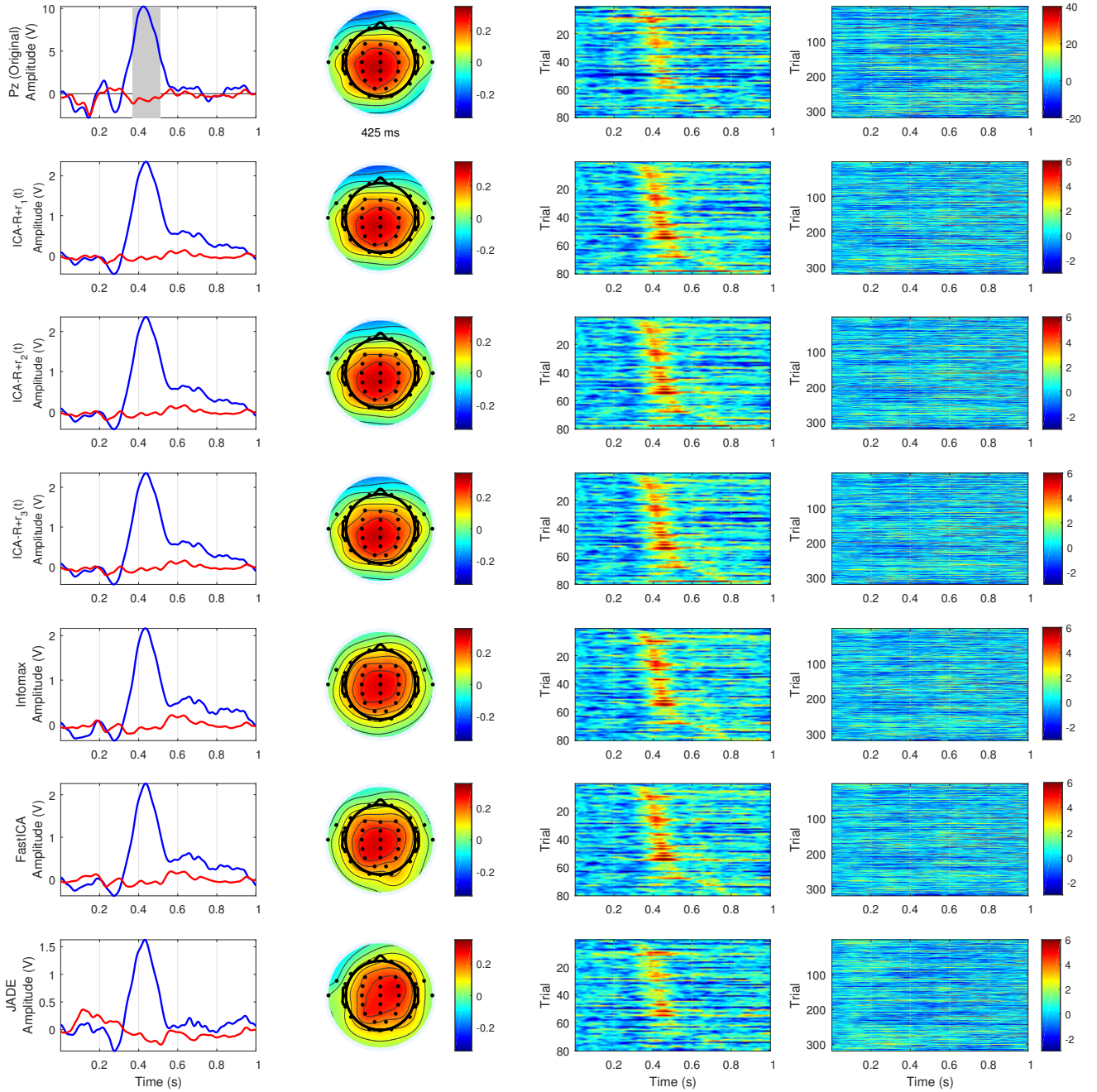


Figure 7: Results obtained from a subject's EEG recording where each row gives the results taken from raw EEG (Channel Pz), one-unit ICA-R with  $r_1(t)$ ,  $r_2(t)$  and  $r_3(t)$ , Infomax ICA, FastICA, and the JADE algorithms. The first, second, third and fourth columns represent respectively the trial-averaged target (blue)/non-target (red) signals, scalp distribution, single-trial target and non-target epochs from different methods. In the first column, the grey region represents the P300 time region identified from trial-averaged target signal and used as  $\mathcal{T}$  for generating  $r_1(t)$ ,  $r_2(t)$  and  $r_3(t)$ . In the second column, the scalp distribution for raw EEG is taken at the P300 peak latency (i.e. 0.425 s) from trial-averaged target signals. In the third column, the single-trial target epoch is sorted based on the ERP peak latencies identified from ICA-R with  $r_2(t)$ . Note that the source signals of the ICA methods cannot reflect the real amplitude because of the scaling ambiguity and the zero-mean and unit-variance source signal assumptions.

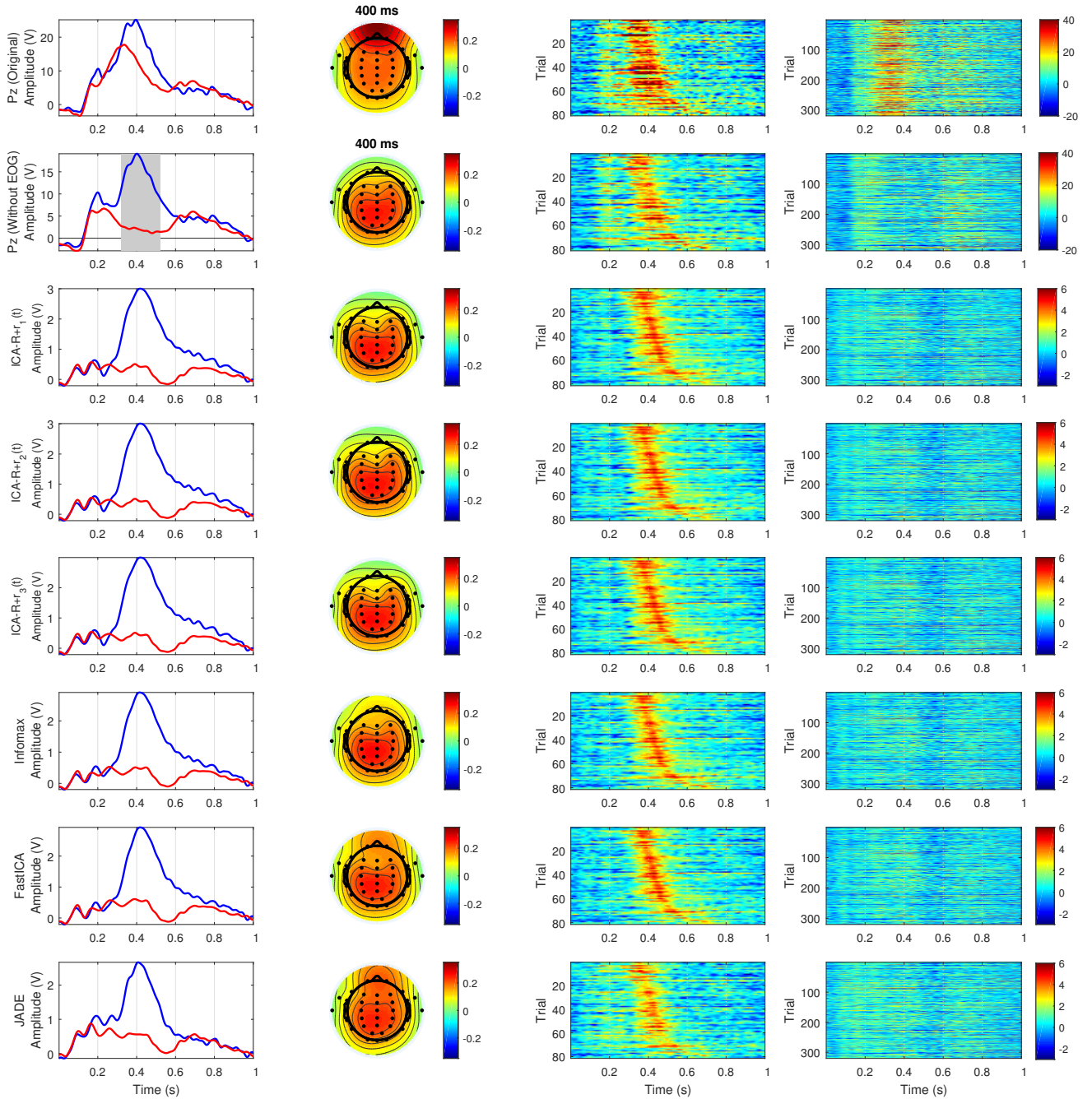


Figure 8: Similar to Fig. 7. Except that, the first and second rows show respectively the target and non-target signals before and after eye-blink removal using the one-unit ICA-R. The third, fourth and fifth rows show respectively the extracted signals when the one-unit ICA-R with  $r_1(t)$ ,  $r_2(t)$  and  $r_3(t)$  were applied on the artifact-free signals in the second row. The sixth, seventh and eighth rows represent the source signal selected from Infomax, FastICA and JADE algorithm respectively.

Thirdly, a reference signal is generated by training a discriminative-based spatial filter. Lastly, the eye-blink artifacts were extracted by using the proposed ICA-R and subsequently removed from the EEG recording by using the equation below:

$$\tilde{\mathbf{x}}(t) = \mathbf{x}(t) - \hat{\mathbf{a}}_e y_e(t) \quad (17)$$

where  $\tilde{\mathbf{x}}(t)$  represents the multi-channel EEG signals that are free of artifacts,  $y_e(t)$  represents the source signal that is belong to the eye-blink activity,  $\hat{\mathbf{a}}$  represents the estimated mixing vector of  $y_e(t)$ .

The second row of Fig. 8 shows the resultant target and non-target signals after the eye-blink removal. It can be observed that the scalp distribution of the trial-averaged target signal at 400 ms resembles the spatio-temporal distribution of a P300 ERP after eye-blink removal. By applying the proposed one-unit ICA-R on the artifact-free trials, the third, fourth and fifth row of Fig. 8 show the source signals extracted by the ICA-R with  $r_1(t)$ ,  $r_2(t)$  and  $r_3(t)$  respectively. Similar to the results in Section 4.3.1, the proposed ICA-R has extracted the P300 ERPs correctly since the extracted signals contain only activity between 300 and 700 ms in target trials. For comparison, the fifth, sixth and seventh rows in Fig. 8 represent the source signals selected from Infomax ICA, FastICA and JADE algorithms respectively.

#### 4.4. General Discussion

Overall, the results in Fig. 7 and 8 indicate that the proposed ICA-R method can extract the desired ERP effectively from the real ERP dataset. The results also show that the signals extracted from our proposed method are almost identical to the source signals extracted from the Infomax ICA. However, it is observed that the FastICA and JADE algorithms may have failed to isolate irrelevant EEG activities and extract P300 ERPs successfully because the scalp distribution of their extracted signals are different from the P300 scalp distribution obtained from raw EEG signals. This is especially true when applying the methods on heavily-contaminated trials as shown in Fig. 8. In comparison to the Infomax ICA, the proposed ICA-R may have better extracted signal because in Fig. 8, the scalp distribution of the source signal from Infomax ICA still contains some activities in the frontal region. By examining the ICA decomposition result of Infomax ICA in Fig. 9, it is observed that there is another source signal that represents the activity from the frontal region and Infomax ICA may have not separate this source signal effectively from the source signal of the desired ERP.

When comparing different reference signals, the finding is that all the reference signals (i.e.  $r_1(t)$ ,  $r_2(t)$  and  $r_3(t)$ ) can guide the proposed one-unit ICA-R to

extract the desired P300 ERP correctly. However, by considering that the estimated latency variation  $\sigma_\tau$  of the real P300 ERP is approximately 75 ms in the above case studies, a reference signal such as  $r_2(t)$  is still recommended since it ensures a more accurate extraction.

Another observation that can be made from the studies is the strength and weakness of the discriminative-based reference signal. Firstly, although the discriminative-based reference signal is developed mainly for the ERP, the concept can be applied to other type of EEG signals such as eye-blinks especially when these signals have a distinct scalp distribution and their activation time can be identified. In fact, the performance of the proposed one-unit ICA-R is potentially better for the signal such as eye-blinks because the time region of eye-blinks can be identified precisely due to their greater signal amplitude. However, as shown in Section 4.3.2, the weakness of the discriminative-based reference signal is also its machine-learning characteristic. Specifically, although it can minimise random noises through time-averaging, it cannot learn the mean scalp distribution of the desired ERP from the training samples that are consistently corrupted by a certain artifact.

In general, the proposed method can be applied directly when there is no excessive ocular artifacts in an EEG recording. However, in cases where the contamination from the ocular artifacts is inevitable, Section 4.3.2 demonstrated how the contaminated EEG recordings can be treated by one-unit ICA-R before being used for extracting the desired ERP. In this paper, it is beyond our scope to investigate the effectiveness of the two-stage extraction method on other types of ocular artifacts such as horizontal and vertical eye movements. But, by considering that the reference signal for these two types of eye movements can be recorded from channel HEOG and VEOG directly, the two-stage extraction method opens doors for investigators to allow more natural viewing of stimuli in ERP studies.

Another interesting question that arises from this real EEG study is the huge performance difference observed across various traditional ICA methods while theoretically, these methods should provide similar solutions. To understand this issue, the ICA decomposition results for Infomax ICA, FastICA and JADE are plotted in Fig. 9, 10 and 11 respectively. Overall, the decomposition results from all traditional ICA methods are different. Since Infomax ICA has greater accuracy, its decomposition result is first examined in Fig. 9. The results indicate that Infomax ICA extracted far more source signals whose scalp distribution reflects activity originated from only one EEG electrode (e.g. component 9, 11, 17, 18,

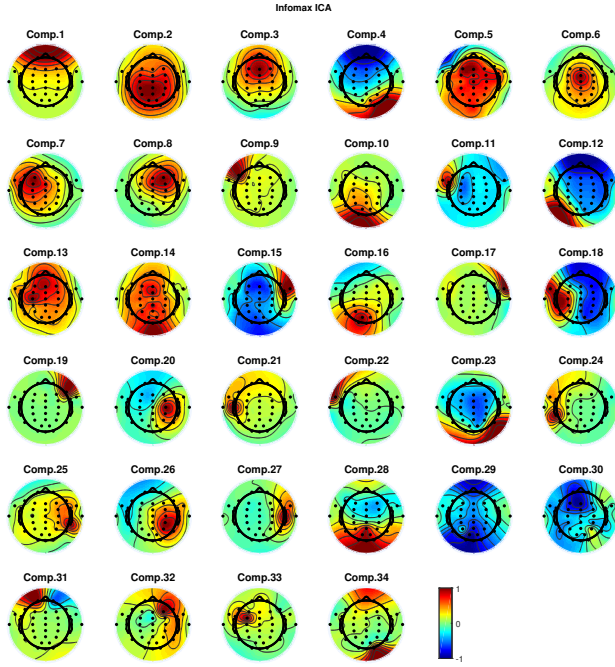


Figure 9: Normalized scalp distribution represented by source signals extracted from Infomax ICA

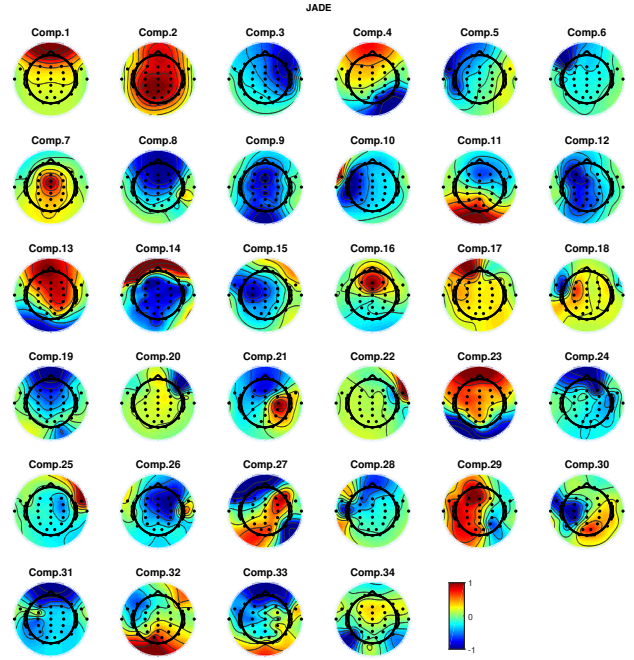


Figure 11: Normalized scalp distribution represented by source signals extracted from JADE

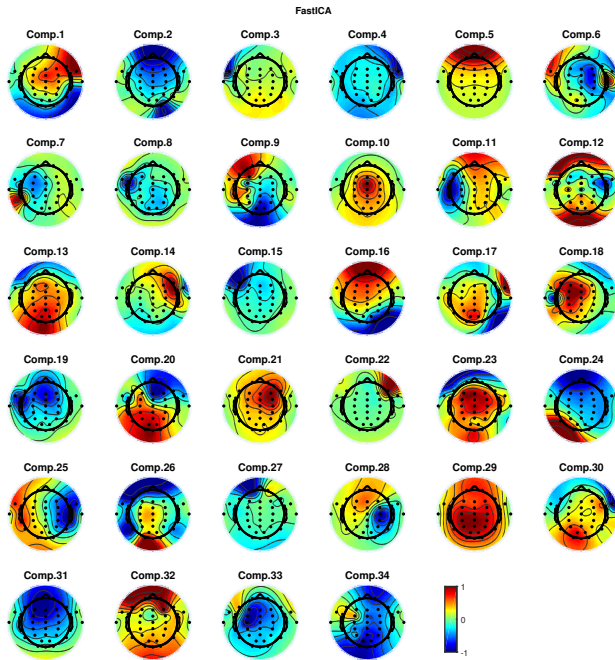


Figure 10: Normalized scalp distribution represented by source signals extracted from FastICA

21, 22, 24, 25, 27, 33) when compared to FastICA and JADE in Fig. 10 and 11 respectively. These extracted source signals are potentially spurious since by volume conduction effect, signal originated within the brain should propagate and reach several scalp electrodes at different voltage levels simultaneously [56]. Thus, the result implies the square-mixing assumption in traditional ICA may not be necessarily true in EEG application. As in this case, the number of sources in a mixed signal could be lower than the number of electrodes. If this observation is true, then FastICA and JADE may have failed to cope with this invalid assumption and thus yielding a poor signal decomposition result as shown in both the semi-simulated and real EEG dataset.

In contrast, the proposed one-unit ICA-R with  $r_2(t)$  may have better performance compared to traditional ICA potentially because it only extracts one source signal. Thus, unlike traditional ICA, the estimation of its demixing vector  $\mathbf{w}$  can be performed without depending on the accuracy of the demixing vectors of the other source signals.

## 5. Conclusion

In this paper, a new approach of one-unit ICA-R has been presented for the problem of ERP extraction at the single-trial level. In cases where ERP-of-interest is known a priori, our proposed one-unit ICA-R relies on the estimated time-region of an ERP to



build a reference signal so that the ICA-R can be guided to extract the source signal of the desired ERP automatically.

Three different ways to generate reference signals were proposed for our one-unit ICA-R. Our simulation shows that although all the reference signals can be used for extracting ERP, the discriminative-based reference signal is preferred because of its ability to reflect the ERP's trial-to-trial variations more accurately. The adaptability of the discriminative-based reference signal helps the ICA-R to achieve a more accurate and consistent extraction performance compared to the other two reference signals.

When compared to the popular Infomax ICA, both the simulation and real study shows that the proposed method is equally efficient in extracting the desired ERP at single-trial level. However, since the proposed method avoids human intervention and uses considerably lower computational time, one-unit ICA-R is a more effective tool for analyzing ERP.

## Acknowledgment

The author would like to specially thank Melissa Black and Lim Yi Huey for editing the manuscript.

## 6. References

- [1] S. J. Luck, *An Introduction to the Event Related Potential Technique*, Cambridge: MIT Press, 2005.
- [2] M. Fabiani, G. Gratton and K. D. Federmeier, "Event-related potential - methods, theory, and applications", in *Handbook of Psychophysiology*, 3rd Ed, J. T. Cacioppo, L. G. Tassinary and G. Berntson Eds. Cambridge: Cambridge University Press, 2007, pp. 85-119.
- [3] C. A. Nelson and J. P. McCleery, "Use of Event-Related Potentials in the Study of Typical and Atypical Development", *J. of the American Academy of Child and Adolescent Psychiatry*, vol. 47, 1252-1261, Nov. 2008.
- [4] C. C. Duncan *et al.*, "Event-related potentials in clinical research: Guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400", *Clinical Neurophysiology*, vol. 120, no. 11, pp. 1883-1908, Nov. 2009.
- [5] J. Polich, "Updating P300: an integrative theory of P3a and P3b", *Clinical Neuro physiology*, vol. 118, pp. 2128-2148, Oct. 2007.
- [6] S. H. Patel and P. N. Azzam, "Characterization of N200 and P300: Selected studies of the event-related potential", *Int. J. Medical Sci.*, vol. 2, pp. 147-154, October 2005.
- [7] B. I. Turetsky, J. Raz and G. G. Fein, "Estimation of trial-to-trial variation in evoked potential signals by smoothing across trials", *Psychophysiology*, vol. 26, no. 6, pp. 700-712, Nov. 1989.
- [8] J. Möcks, T. Gasser, D. T. Pham and W. Köhler, "Trial-to-trial variability of single potentials - methodological concepts and results", *Int. J. Neuroscience*, vol. 33, pp. 25-32, Mar. 1987.
- [9] D. Ravedn and J. Polich, "On P300 measurement stability: habituation, intra-trial block variation, and ultradian rhythms", *Biological Psychology*, vol. 51, pp. 59-76, Oct. 1999.
- [10] R. Q. Quiroga and H. Garcia, "Single-trial event-related potentials with wavelet denoising", *Clinical Neurophysiology*, vol. 114, pp. 376-390, Feb. 2003.
- [11] Z. Wang, A. Maier, D. A. Leopold, N. K. Logothetis and H. Liang, "Single-trial evoked potential estimation using wavelets", *Computers in Biology and Medicine*, vol. 37, pp. 463-473, April 2007.
- [12] W. Truccolo, K. H. Knuth, A. Shah, S. L. Bressler, C. E. Schroeder and M. Z. Ding, "Estimation of single-trial multicomponent ERPs: Differentially variable component analysis (dVCA)", *Biological Cybern.*, vol. 89, no. 6, pp. 426-438, Dec. 2003.
- [13] L. Xu, P. Stoica, J. Li, S. L. Bressler, X. Shao and M. Ding, "ASEO: A method for the simultaneous estimation of single-trial event-related potentials and ongoing brain activities", *IEEE Trans. Biomed. Eng.*, vol. 56, pp. 111-121, Jan. 2009.
- [14] P. A. Karjalainen, J. P. Kaipio, A. S. Koistinen and M. Vauhkonen, "Subspace regularization method for the single-trial estimation of evoked potentials", *IEEE Trans. Biomed. Eng.*, vol. 46, July 1999.
- [15] P. O. Ranta-aho, A. S. Koistinen, J. O. Ollikainen, J. P. Kaipio, J. Partanen and P. A. Karjalainen, "Single-trial estimation of multichannel evoked-potential measurements", *IEEE Trans. Biomed. Eng.*, vol. 50, pp. 189-196, Feb. 2003.
- [16] R. Li, J. C. Principe, M. Bradley and V. Ferrari, "A spatiotemporal filtering methodology for single-trial ERP component estimation", *IEEE Trans. Biomed. Eng.*, vol. 56, pp. 83-92, Jan. 2009.
- [17] D. Jarchi, S. Sanei, J. C. Principe and B. Makkiabadi, "A new spatiotemporal filtering method for single-trial estimation of correlated ERP subcomponents", vol. 58, pp. 132-143, Jan. 2011.
- [18] M. Zibulevsky and Y. Y. Zeevi, "Extraction of a source from multichannel data using sparse decomposition", *Neurocomputing*, vol. 49, pp. 163-173, 2002.
- [19] R. Vigário and E. Oja, "BSS and ICA in neuroinformatics: From current practices to open challenges", *IEEE Reviews in Biomed. Eng.*, vol. 1, pp. 50-61, 2008.
- [20] T. P. Jung, S. Makeig, M. Westerfield, J. Townsend, E. Courchesne and T. J. Sejnowski, "Analysis and visualization of single-trial event-related potentials", *Hum. Brain Mapp.* vol. 14, no. 3, pp. 166-185, Nov 2001.
- [21] S. Makeig, T. P. Jung, A. J. Bell, D. Ghahremani and T. J. Sejnowski, "Blind separation of auditory event-related brain responses into independent component", *Proc. Natl. Acad. Sci.*, vol. 94, pp. 20979-10984, Sept. 1997.
- [22] M. De Vos, J. D. Thorne, G. Yovel and S. Debener, "Let's face it, from trial to trial: comparing procedures for N170 single-trial estimation", *Neuroimage*, vol. 15, pp. 1196-1202, Aug. 2012.
- [23] N. Xu, X. Gao, B. Hong, X. Miao, S. Gao and F. Yang, "BCI competition 2003-data set IIb: enhancing P300 wave detection using ICA-based subspace projections for BCI applications", *IEEE Trans. Biomed. Eng.*, vol. 51, pp. 1067-1072, June 2004.
- [24] H. Serby, E. Yom-Tov and G. F. Inbar, "An improved P300-based brain-computer interface", *IEEE Trans. Neural Syst. and Rehabil. Eng.*, vol. 13, pp. 89-98, March 2005.
- [25] F. Cong, I. Kalyakin, T. Huttunen-Scott, H. Li, H. Lyytinen and T. Ristaniemi, "Single-trial based independent component analysis on mismatch negativity in children", *Int. J. Neural Syst.*, vol. 20, pp. 279-292, Aug. 2010.

- [26] J. Onton, M. Westerfield, J. Townsend and S. Makeig, "Imaging human EEG dynamics using independent component analysis", *Neuroscience and Biobehavioral Reviews*, vol. 30, pp. 808-822, 2006.
- [27] D. Iyer and G. Zouridakis, "Single-trial evoked potential estimation: comparison between independent component analysis and wavelet denoising", *Clinical Neurophysiology*, vol. 118, pp. 495-504, March 2007.
- [28] C. W. Hesse and C. J. James, "On semi-blind source separation using spatial constraints with applications in EEG analysis", *IEEE Trans. Biomed. Eng.*, vol. 53, pp. 2525-2534, Dec. 2006.
- [29] M. De Vos, L. D. Lathauwer and S. V. Huffel, "Spatially constrained ICA algorithm with an application in EEG processing", *Signal Processing*, vol. 91, pp. 1963-1972, 2011.
- [30] C. Porcaro, D. Ostwald and A. P. Bagshaw, "Functional source separation improves the quality of single-trial visual evoked potentials recorded during concurrent EEG-fMRI", *NeuroImage*, vol. 50, pp. 112-123, 2010.
- [31] W. Lu and J.C. Rajapakse, "ICA with reference", in *Proc. 3rd Int. Conf. Independent Component Analysis and Blind Signal Separation: ICA2001*, San Diego, pp. 120-125.
- [32] C. J. James and O. J. Gibson, "Temporally constrained ICA: An application to artifact rejection in electromagnetic brain signal analysis", *IEEE Trans. Biomed. Eng.*, vol. 50, pp. 1108-1115, Sept. 2003.
- [33] Z. L. Zhang, "Morphologically constrained ICA for extracting weak temporally correlated", *Neurocomputing*, vol. 71, pp. 1669-1679, March 2008.
- [34] Q. H. Lin, Y. R. Zheng, F. L. Yin, H. Liang and V. D. Calhoun, "A fast algorithm for one-unit ICA-R", *Information Sciences*, vol. 177, pp. 1265-1275, Sep. 2006.
- [35] A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications", *Neural Networks*, vol. 13, pp. 411-430, June 2000.
- [36] C. J. James and C. W. Hesse, "Independent component analysis for biomedical signals", *Physiological Measurement*, vol. 26, pp. R15-R39, Feb. 2005.
- [37] A. Hyvärinen, "One-unit contrast functions for independent component analysis: a statistical analysis", In *Proc. 1997 IEEE Workshop on Neural Networks for Signal Process.*, Amelia Island, Sept. 1997, pp. 388-397.
- [38] V. Zarzoso and A. Hyvärinen, "Iterative algorithms", in *Handbook of Blind Source Separation*, P. Comon and C. Jutten Eds. Oxford: Academic Press, 2010, ch. 6, sec. 6.9.2, pp. 207.
- [39] A. Cichocki and S. Amari, "Sequential Blind Signal Extraction", in *Adaptive blind signal and image Processing*. West Sussex, England: John Wiley & Sons, 2002, ch. 5, sec. 5.2.5, pp. 191-193.
- [40] S. Scholler, S. Bosse, M. S. Treder, B. Blankertz, G. Curio, K. R. Müller and T. Wiegand *et al.*, "Toward a direct measure of video quality perception using EEG", *IEEE Trans. Image Process.*, vol. 21, no. 5, pp. 2619-2629, May 2012.
- [41] B. Blankertz, S. Lemm, M. Treder, S. Haufe and K-R Müller, "Single-trial analysis and classification of ERP components - A tutorial", *Neuroimage*, vol. 56, pp. 814-825, Jun 2010.
- [42] W. L. Lee, T. Tan, Torbjörn Falkmer and Y. H. Leung, "Single-trial multi-channel N170 estimation using linear discriminant analysis (LDA)", in *Proc. Int. Conf. on Neural Information Processing, ICONIP 2012*, Doha, Nov. 2012, pp. 347-355.
- [43] B. Blankertz *et al.*, "The BCI competition 2003: Progress and perspectives in detection and discrimination of EEG single trials", *IEEE. Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1044-1051, Jun. 2004.
- [44] T. P. Jung, S. Makeig, M. Westerfield, J. Townsend, E. Courchesne and T. J. Sejnowski, "Analysis and visualization of single-trial event-related potentials", *Human Brain Mapping* vol. 14, no. 3, pp. 166-185, Nov. 2001.
- [45] T. W. Lee, M. Girolami and T. J. Sejnowski, "Independent component analysis using an extended Infomax algorithm for mixed subgaussian and supergaussian sources", *Neural Computation*, vol. 11, pp. 417-441, 1999.
- [46] J. F. Cardoso and A. Souloumiac, "Blind beamforming for non-Gaussian signals", *IEE Proceedings F. (Radar and Signal Processing)*, vol. 140, pp. 362-370, 1993.
- [47] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis", *J. Neuroscience Methods*, vol. 134, no. 1, 9-21, Mar. 2004.
- [48] FastICA package for MATLAB [Online]. Available: <http://research.ics.aalto.fi/ica/fastica/>.
- [49] W. Lu and J. C. Rajapakse, "ICA with Reference", in *Neurocomputing*, vol. 69, pp. 2244-2257, Feb. 2006.
- [50] S. Ting, T. Tan, G. West, A. Squelch and J. Foster, "Quantitative assessment of 2D versus 3D visualization modalities", in *Visual Communications and Image Processing*, Tainan, pp. 6-9, Nov. 2011.
- [51] T. Tan, "Brain responses to 2d versus 3d images using EEG". Available: <http://researchdata.and.s.org.au/brain-responses-to-2d-versus-3d-images-using-eeeg>
- [52] M. M. Murray, D. Brunet and C. M. Michel, "Topographic ERP Analyses: A Step-by-Step Tutorial Review", *Brain Topography*, vol. 20, no. 4, pp. 249-264, Mar. 2008.
- [53] R. D. Pascual-Marqui, C. M. Michel and D. Lehmann, "Segmentation of Brain Electrical Activity into Microstates: Model Estimation and Validation", *IEEE Trans. Biomed. Eng.*, vol. 42, no. 7, pp. 658-665, Jul. 1995.
- [54] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, 4th ed., Burlington: Academic Press, 2008.
- [55] W. L. Lee, T. Tan and Y. H. Leung, "An improved P300 extraction using ICA-R for P300-BCI speller", in *Conf. Proc. of IEEE Eng. Med. Biol. Soc. 2013, Osaka, 2013*, pp. 7064-7067.
- [56] P. L. Nunez and R. Srinivasan, *Electric fields of the brain: Neurophysics of EEG*, Oxford: Oxford University Press, 2006.