Accepted version of an article published in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 27 (9), pp. 1883 - 1892
DOI: https://doi.org/10.1109/TNSRE.2019.2934645

Asynchronous Control of ERP-based BCI Spellers Using Steady-State Visual Evoked Potentials Elicited by Peripheral Stimuli

Eduardo Santamaría-Vázquez, Víctor Martínez-Cagigal, Javier Gomez-Pilar, and Roberto Hornero, *Senior Member, IEEE*

***Abstract*—Brain–computer interface (BCI) spellers based on event related potentials (ERPs) are intrinsically synchronous systems. Therefore, selections are constantly made, even when users are not paying attention to the stimuli. This poses a major limitation in real-life applications, in which an asynchronous control is required. The aim of this study is to design, develop and test a novel method to discriminate whether the user is controlling the system (i.e., control state) or is engaged in other task (i.e., non-control state). To achieve such asynchronous control, our method detects the steady-state visual evoked potentials (SSVEPs) elicited by peripheral stimuli of ERP-based spellers. A characterization experiment was conducted to investigate several aspects of this phenomenon. Then, the proposed method was validated in offline and online sessions. A total of 20 healthy subjects participated the experiments. The proposed method achieved an average accuracy of 95.5% for control state detection during the online sessions, providing a reliable asynchronous control. Furthermore, our approach is independent of the ERP classification stage, and to the best of our knowledge, is the first procedure that does not need to extend the duration of the calibration sessions to acquire non-control observations.**

***Index Terms*—Brain–computer interfaces, event-related potentials, asynchrony, control-state detection, steady-state visual evoked potentials, P300.**

I. Introduction

BRAIN–computer interfaces (BCIs) provide a direct pathway between the brain and an external device. These systems use the neural activity to identify the user's intentions and translate them into commands [1], [2]. The main application of BCI systems is to create a new channel of communication in real time for severely disabled people. BCI systems enhance or restore their capability to relate to their environment [1]. Electroencephalography (EEG) is the most common technique to register the neural activity in BCI systems [1]. EEG uses electrodes to register the electrical signal generated by the superficial neurons close to them. This signal is the superposition of rhythmic and transient components that reflect the underlying activity of the brain. Particularly, event-related potentials (ERPs) are the natural response of the brain to different types of events [3]. ERPs elicited by a single visual stimulus (e.g., a brief flash) are known as visual ERPs [3].

The first ERP-based BCI was proposed by Farwell and Donchin [4]. The system used the row-column paradigm (RCP) to determine the intentions of the user. The RCP displays a matrix (i.e., speller) with several commands. The rows and columns of the matrix are randomly highlighted a predefined number of times (i.e., sequences). The user has to stare at the desired command and count the flashes (i.e., target stimuli), while ignoring the other stimuli (i.e., non-target stimuli). Target stimuli elicits visual ERPs in the EEG just after each flash. The signal processing stage detects the visual ERPs and determines the command. This BCI is also named as P300-based speller because of the P300 potential, which is the most prominent peak of visual ERPs elicited by target stimuli [1], [3]. Recent research has improved the accuracy and the speed of this BCI through new stimulation paradigms or signal processing methods [5]–[7]. Nevertheless, ERP-based spellers are synchronous systems. They always make a selection for every trial, even when the user is not paying attention to the stimuli. This poses a major limitation for most applications, such as wheelchair control or web browsers, where an asynchronous approach should be a key feature [8], [9]. For this reason, detecting whether the user wants to make a selection with the speller (i.e., control state) or is engaged in other task (i.e., non-control state) is a crucial issue to bridge the gap between laboratory and real-life BCI applications.

In recent years, several studies have addressed the asynchronous control in ERP-based spellers following two main strategies: (i) algorithms that define a threshold dependent on the output scores of the ERP classification stage [6], [8]–[12]; or (ii) hybrid BCIs that combine different control signals [13]–[15]. Nevertheless, these approaches present a number of drawbacks. For instance, algorithms that rely on the output scores of the ERP classification stage have high inter-session variability [5]. Small changes in the amplitude or latency of the ERPs could override the threshold and cause a drastic decrease in the peak accuracy of these methods. In fact, thresholds must be recalibrated before each session with the BCI system. This procedure is time consuming, affects the usability of the system and is frustrating for users [5],

This study was partially funded by projects DPI2017-84280-R of 'Ministerio de Ciencia, Innovación y Universidades' and 'European Regional Development Fund' (FEDER), and the project 'Análisis y correlación entre el genoma completo y la actividad cerebral para la ayuda en el diagnóstico de la enfermedad de Alzheimer' (Inter-regional cooperation program VA Spain-Portugal POCTEP 2014-2020) of the European Commission and FEDER. V. Martínez-Cagigal was in receipt of a PIF-UVa grant of the University of Valladolid.

The authors are with the Biomedical Engineering Group, E.T.S.I Telecomunicación, University of Valladolid, Paseo de Belén 15, 47011, Valladolid, Spain (e-mail: eduardo.santamaria@gib.tel.uva.es; victor.martinez@gib.tel.uva.es; javier.gomez@gib.tel.uva.es; robhor@tel.uva.es)

[10], [12]. Moreover, previous approaches require to record additional data of the non-control state of the users (i.e., non-control trials) to calculate the threshold, increasing the duration of the calibration sessions. On the other hand, hybrid BCIs increase the complexity of the system, requiring two or more different stimulation interfaces. In these BCIs, the user has to be able to manage with two or more control signals, making them more demanding. Consequently, hybrid BCIs might be a major challenge for certain users who cannot maintain high levels of concentration [6], [16]. Algorithms independent of the ERP classification stage may help to overcome these limitations. In this regard, Pinegger et al. [17] proposed a method based on the hypothesis that the flashing frequency of the RCP should be represented somehow in the EEG, reaching an average accuracy of 79.5% for control state detection. However, this method still needs to double the duration of the calibration sessions in order to acquire non-control trials.

The aim of this study is to design, develop and test a novel method for control state detection able to overcome the limitations of previous approaches. We hypothesize that peripheral stimuli (i.e., non-target stimuli) of ERP-based spellers elicit a weak steady-state visual evoked potential (SSVEP). SSVEPs are waveforms, similar to a sinusoid, that appear with repetitive visual stimuli i the EEG [3], [18]. In order to test the hypothesis, we performed a characterization experiment to study the cause of this phenomenon and its dependence on the stimulation frequency. Finally, we designed and developed a novel method for asynchronous control based on the detection of the SSVEPs elicited by peripheral stimuli. Our approach is independent of the ERP classification stage and does not use additional control signals. Furthermore, to the best of our knowledge, it is the first asynchrony algorithm that does not need to extend the duration of the calibration sessions to acquire non-control trials. The efficacy of the method was tested in offline and online experiments.

II. Methods

A. Signal acquisition and subjects

This study consist of 3 tests: characterization, offline and online experiments. For the characterization experiment, EEG signal was recorded using 16 active electrodes: F3, Fz, F4, C3, Cz, C4, CPz, P3, Pz, P4, PO3, PO7, POz, PO4, PO8 and Oz, according to the extended International 10–20 System distribution. Two electrodes in FPz and the earlobe were used as ground and reference, respectively. This distribution was chosen to favor the detection of the SSVEPs [18]. For the offline and online experiments, EEG signals were recorded using 8 active electrodes, placed at Fz, Cz, Pz, P3, P4, PO7, PO8 and Oz. This configuration was not only chosen based on the results of the characterization experiment, but also because it is commonly used in ERP-based spellers [9], [10]. A g.USBamp (g.Tec, *Guger Technologies*, Austria) was used to amplify and convert the signal into the digital domain with a sampling frequency of 256 Hz. A novel BCI platform, MEDUSA®, was developed and used to present the stimuli, record and save the data and process the signal during the online sessions [19].

Twenty healthy subjects participated in the experiments, divided into two groups. Five subjects (mean age: 31.2 ± 4.8 years; 4 males; 1 female) took part in the characterization experiment. The remaining 15 subjects (mean age: 26.1 ± 2.3 years; 11 males; 4 females) took part in the offline and online sessions. The experimental protocol was approved by the local ethics committee and all participants gave their informed consent.

B. Characterization experiment

The characterization experiment was aimed at analizing the SSVEPs elicited by RCP. The stimulation frequency of ERP-based spellers is reflected in the EEG and can be used to determine the control state in an asynchronous BCI system [17]. However, the origin and characteristics of this phenomenon have not been studied yet. We hypothesize that non-target stimuli of RCP trigger a weak SSVEP in the user's EEG. Based on this assumption, two analysis were conducted:

1) First analysis: This analysis was designed to investigate how the characteristics of the SSVEP vary as a function of the stimulation frequency. This frequency is the inverse of the stimulus onset asynchrony (SOA), which is the time between two consecutive flashes calculated as the sum of the stimulus duration and inter-stimulus time. Therefore, the stimulation frequency is defined as $f_{\text{st}} = 1/\text{SOA}$. Participants had to spell 6 words (i.e., runs) of 6 letters (i.e., trials) using the matrix shown in Fig. 1a. In one trial, each row and column was highlighted 15 times (i.e., sequences). Each run had a different stimulation frequency as shown in Table I.

2) Second analysis: This analysis was designed to find out the cause that provokes the SSVEPs. Four runs were performed with different stimulation matrices. As previously, each run had 6 trials of 15 sequences. SOA was fixed to 175 ms ($f_{\text{st}} = 5.71$Hz) [17]. For the first run, the overt matrix shown in Fig. 1b was used. Participants were asked to stare at the black space while the rows and columns were highlighted. Therefore, participants only saw stimuli with their peripheral visual field. For the second run, the covert matrix shown in Fig. 1c was used. This matrix only had one letter in the centre. Participants were asked to fix the gaze in this letter while it was randomly highlighted. Unlike in the first run, participants only saw stimuli in the central region of their visual field. For the last two runs, the RCP matrix shown in Fig. 1a was used. These runs consisted of one control run and one non-control run. In the control run, participants were asked to spell 6 characters. In the non-control run, participants were watching a video without attending the stimuli. This analysis was intended to study whether the SSVEPs are provoked by the peripheral stimuli of the RCP. In that case, the EEG signal in the overt mode should only show the SSVEP, but no ERPs with the P300 component would be found. Conversely, the EEG signal of the covert mode should only contain ERPs, but no SSVEP. Accordingly, the control mode would show the superposition of the SSVEP and ERPs, while in the non-control mode none of these waveforms should be found.

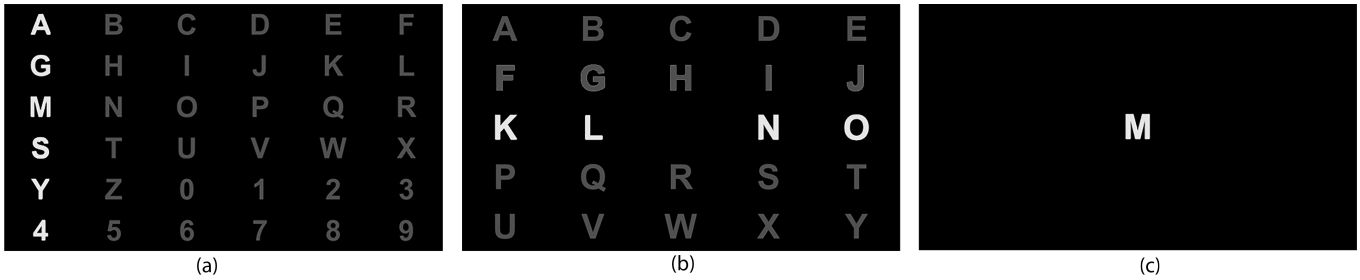

Fig. 1. Matrices used during the experiments: (a) RCP matrix, (b) overt matrix, (c) covert matrix.

TABLE I
STIMULATION PARAMETERS FOR THE CHARACTERIZATION EXPERIMENT

Run	SOA	Inter-stimulus time	Stimulus duration	Flashing frequency
1	475 ms	400 ms	75 ms	2.10 Hz
2	375 ms	300 ms	75 ms	2.67 Hz
3	275 ms	200 ms	75 ms	3.64 Hz
4	175 ms	100 ms	75 ms	5.71 Hz
5	150 ms	75 ms	75 ms	6.66 Hz
6	125 ms	50 ms	75 ms	8.00 Hz

SOA: stimulus onset asynchrony

C. Proposed method for control state detection

The Oddball Steady Response Detection (OSRD) method is our novel algorithm for detecting the user's control state in ERP-based spellers. OSRD provides a binary output $y \in \{0, 1\}$ for each trial. When control state is detected ($y = 1$), the system selects a command. Conversely, when non-control state is detected ($y = 0$), no further actions are taken. OSRD takes the EEG signal corresponding to one trial as input (i.e., signal from the first stimulus of the first sequence to the last stimulus of the last sequence considered), and performs in real time the following stages:

1) Pre-processing: The objective of this stage is to increase the signal-to-noise ratio. Finite impulse response (FIR) band-pass filter is applied between $[f_{\text{st}} - bw_1/2,\ f_{\text{st}} + bw_1/2]$ Hz, where bw_1 is heuristically set to 2 Hz. The stimulation frequency (f_{st}) is calculated as the difference between onsets of consecutive stimuli. Common Average Reference (CAR) is applied to remove common noisy components in the EEG [10].

2) Feature extraction: The proposed method extracts two features for each trial, which are based on canonical correlation analysis (CCA) and power spectral density (PSD). CCA is a multivariate statistical method that finds underlying correlations between two multidimensional sets of data (i.e., $\boldsymbol{X}$,$\boldsymbol{Y}$) [20]. CCA finds the optimal linear combination $\boldsymbol{x} = \boldsymbol{w}_x^T * \boldsymbol{X}$, $\boldsymbol{y} = \boldsymbol{w}_y^T * \boldsymbol{Y}$ which maximizes the correlation ρ between $\boldsymbol{X}$ and $\boldsymbol{Y}$. In this case, $\boldsymbol{X}$ is the EEG signal of one trial with dimensions $N \times C$, where N is the length of the trial and C is the number of channels. On the other hand, $\boldsymbol{Y}$ is the reference signal, which is a sine wave of frequency f_{st} with dimensions $N \times 1$. An schematic representation of CCA is depicted in Fig. 2a. CCA has been successfully applied in BCIs for the detection of SSVEPs [21], [22]. However, to the best of our knowledge, it has not been applied to asynchronous ERP-based spellers yet. Hence, the first feature (x_1) is the maximum correlation coefficient between the trial signal and the reference:

$$x_1 = \texttt{CCA}(\boldsymbol{X}, \boldsymbol{Y}) \tag{1}$$

The second feature (x_2) is directly derived from the spectrum estimation. Firstly, all channels of the trial are concatenated in a single vector of dimensions $(N \cdot C) \times 1$. Secondly, the PSD of this vector is estimated using the Welch's method. The concatenation of the channels allows increasing the spectral resolution. Afterward, x_2 is calculated as the difference between the mean value of the PSD in a narrow range and the mean value in a wide range (see Fig. 2b), formally:

$$x_2 = \frac{1}{bw_2} \int_{f_{\text{st}}-bw_2/2}^{f_{\text{st}}+bw_2/2} S(f)df - \frac{1}{bw_1} \int_{f_{\text{st}}-bw_1/2}^{f_{\text{st}}+bw_1/2} S(f)df, \tag{2}$$

where $S(f)$ is the PSD, and bw_1 and bw_2 are fixed to 2 Hz and 0.1 Hz, respectively. In practice, these values are chosen to have enough points of the PSD within these bands, making the method more robust against noise.

3) Synthetic non-control observations: ERP-based spellers should be recalibrated frequently owing to the high variability of ERPs [5]. In fact, users have to spell several trials before each BCI session to assure maximum performance. This procedure is time consuming and frustrating for users [5]. Moreover, previous asynchronous ERP-based spellers needed to extend the duration of calibration sessions to acquire non-control trials. Here, we introduce a new approach in order to create synthetic non-control observations from control trials. Consequently, there is no need to register non-control trials in calibration sessions, which is a great advantage in real-life applications. The aforementioned features (i.e., x_1 and x_2) are calculated based on the stimulation frequency, f_{st}. Assuming that the PSD from the EEG maintains its characteristics in a narrow band around f_{st}, non-control observations can be simulated by shifting the stimulation frequency a fixed value f_0 (see Fig. 2c): $f'_{\text{st}} = f_{\text{st}} + f_0$, where f_0 is set to 0.5 Hz. Afterward, this value is used in Eq. 1 and Eq. 2 to calculate x_1 and x_2. This procedure is based on the assumption that the characteristics of the PSD at f'_{st} would be similar to non-control trials. Hence, this approach uses control trials to create synthetic non-control observations.

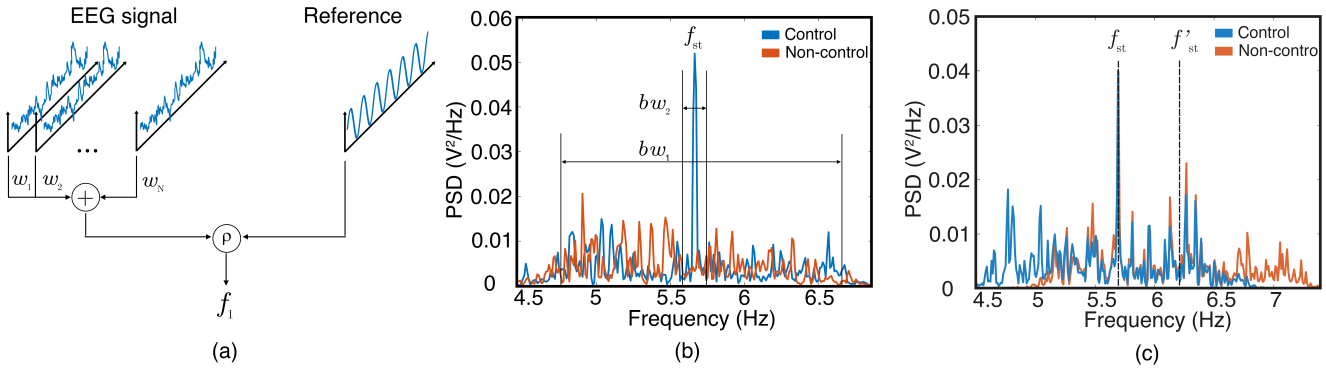

Fig. 2. (a) Schematic representation of the Canonical Correlation Analysis (CCA) used to extract the first feature x_1, where EEG signal represent one trial ($N \times C$) and the reference signal is an ideal sinusoid $y = \sin(2\pi \cdot f_{st})$ ($N \times 1$). CCA calculates the maximum correlation, ρ, between the two datasets. (b) Power spectral density (PSD) of one control and one non-control trials. Parameters bw_1 and bw_2 are used to calculate the second feature x_2. (c) Shifting of the stimulation frequency f_{st} to calculate a synthetic non-control observation from one control trial.

4) Feature classification: The last stage determines the control state for each observation ($y \in \{0, 1\}$). Linear discriminant analysis (LDA) is a linear classifier that applies dimensionality reduction by projecting the data to simultaneously minimize the within-class covariance and maximize the between-class covariance [23]. In this study, LDA is applied due to its extensive use in BCI systems [9], [10], [12], [17]. Previously, input features are normalized using z-score.

D. Offline experiment

The offline experiment was aimed at: (i) analizing the performance of OSRD method, (ii) validating the synthetic non-control observations approach, and (iii) acquiring data to train OSRD method for the online session.

Fifteen participants performed 10 control and 10 non-control runs in 2 sessions. Each run had 6 trials of 15 sequences. In control runs, participants were asked to spell 6 random characters using the matrix shown in Fig. 1a. In non-control runs, participants had to ignore the stimuli while watching a video or reading a text. Therefore, the dataset was composed of 60 control trials and 60 non-control trials per participant.

Two analysis were performed to assess the performance of OSRD method and validate the synthetic non-control observations approach. In the first analysis, leave-one-out (LOO) procedure was applied in the whole dataset for each user, using control and non-control trials (i.e., 120 trials). In each iteration, the classifier was trained using complete trials (i.e., 15 sequences). However, testing features were calculated for each number of sequences. In the second analysis, the evaluation procedure was modified as follows: The training features were calculated only with the control trials, creating from them synthetic non-control observations. When LOO procedure leaved one synthetic non-control observation out for testing, it was replaced by one real non-control observation. Therefore, OSRD was trained with synthetic non-control observations but tested with real non-control observations.

E. Online experiment

The online experiment was aimed at assessing the efficacy of OSRD in a real setting. To this end, the same participants of the offline experiments were asked to spell 4 random words of 6 letters (i.e., 24 trials) using the matrix shown in Fig. 1a. During the first 3 trials of each word, participants had to attend the stimuli. Conversely, during the last 3 trials, participants had to ignore the stimuli while reading a text. We determined the optimal number of sequences for each user as the minimum to reach a training accuracy of 95% in command selection. Whether a participant did not meet this criteria, it was established to 15 sequences.

The control state was determined for each trial applying OSRD method. Afterward, the command was selected by detecting ERPs in the epochs of signal after each stimulus of the RCP paradigm. It should be noted that the system only selected a command whether control state was determined previously. The command selection algorithm had 4 stages. In the preprocessing stage, frequency (i.e., band-pass, 0.5-30 Hz, FIR) and spatial filtering (i.e., CAR) were applied. In the feature extraction stage, subsampling (i.e., 20 Hz) over signal epochs of 800 ms from the stimulus onset was applied. The features of each channel were concatenated to make an observation. In the feature selection stage, backward and forward step-wise regression was applied to select the 60 most relevant features for each user. These features were fed to an LDA to detect the ERPs in the epochs. Each epoch was labelled with the code of the row or column that was highlighted. Finally, the algorithm selected the command corresponding to the row and column that reached the highest score in the classification stage. Further information about this methodology can be found in [10], [12]. Control state detection (i.e., OSRD) and command selection stages were calibrated using the control trials of the offline dataset (i.e., 60 trials). Hence, the training set of the OSRD method was composed by 60 real control observations and 60 synthetic non-control observations.

III. Results

Results of the characterization experiment are shown in Fig. 3 and Fig. 4. The former depicts the SSVEPs for each stimulation frequency of the first characterization analysis. Graphs on the left show the average PSD of all trials, channels

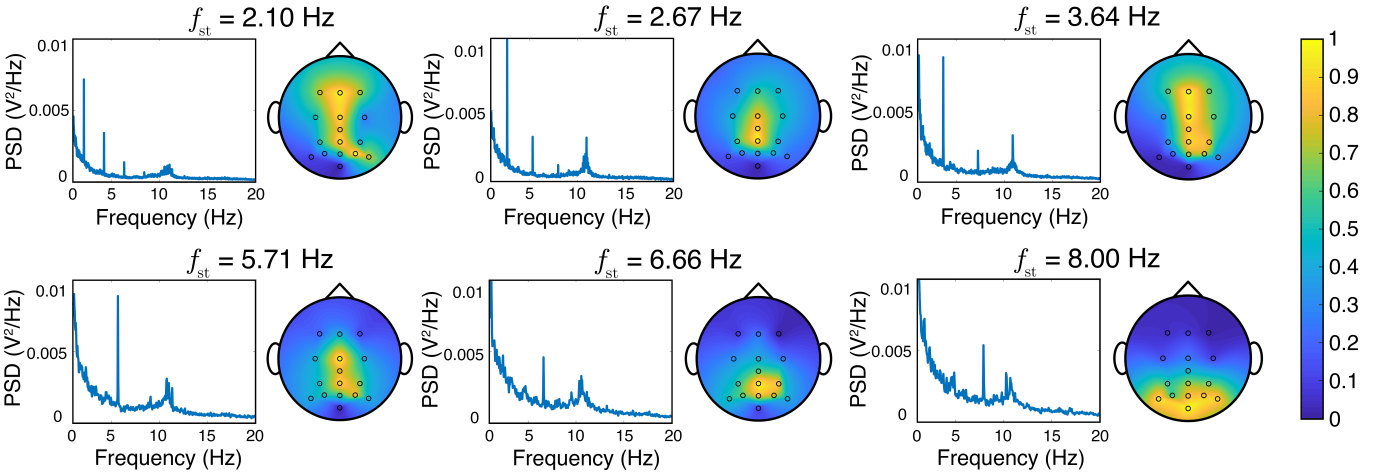

Fig. 3. SSVEPs for different stimulation frequencies. Graphs on the left show the grand average of the PSD for all trials, channels and participants. Topographic plots on the right show the normalized peak value of the PSD at the stimulation frequency averaged for all trials and participants.

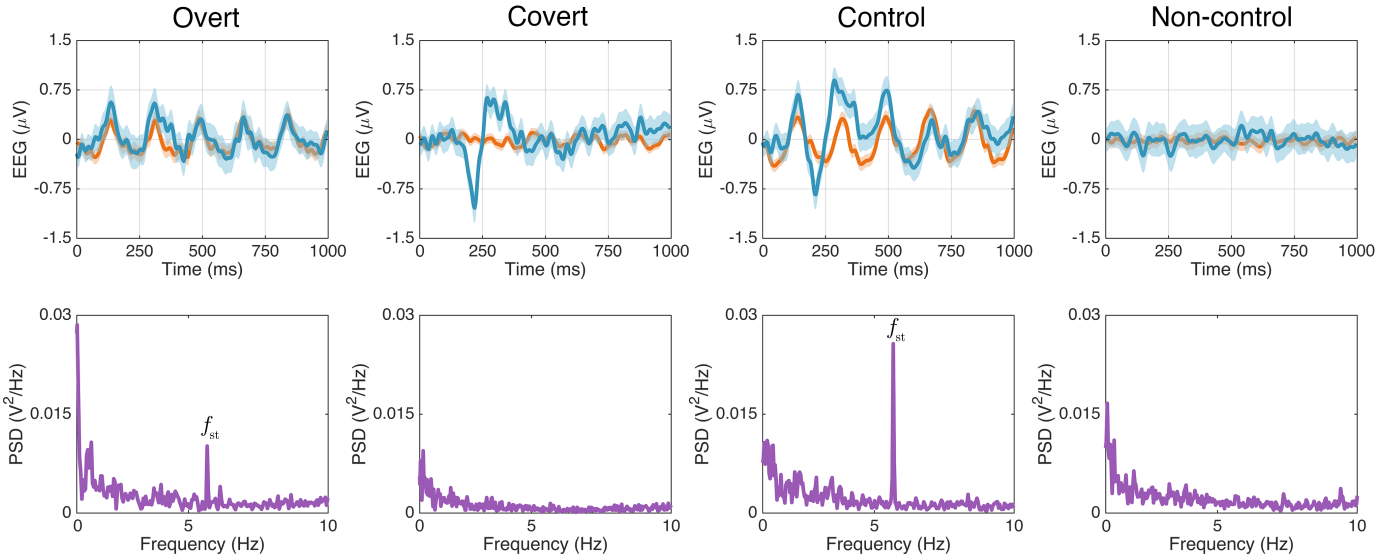

Fig. 4. Temporal and spectral representation of the ERPs in overt, covert, control and non-control modes for participant S01. The upper part of the figure shows the averaged epochs (i.e., 1000 ms after stimuli) in channel Cz. The shaded area represents the 95% confidence interval. The bottom part of the figure shows the averaged PSD in channel Cz. The stimulation frequency in overt and control modes was 5.71 Hz.

and participants. Topographic plots on the right show the normalized power of the SSVEPs per channel (i.e., peak value of the PSD at the stimulation frequency) averaged for all trials and participants. Fig. 4 presents the differences in the EEG signal among the 4 different settings of the second analysis (i.e., overt, covert, control, non-control) for participant S01 in channel Cz. The upper part of the figure shows the averaged epochs of 1000 ms from the stimulus onset. The bottom part of the figure shows the average PSD.

Results of the offline experiments for control state detection are presented in Table II. Test accuracies achieved by OSRD in the two performed analysis are shown for 1, 5, 10 and 15 sequences. The SOA was set to 175 ms (f_{st} = 5.71Hz). This value was chosen to maintain a balance between speed of selection and the SSVEP power (see Fig. 3).

Results of the online experiments are summarized in Table III. The first column details the number of sequences (N_s) used during the session. The results of the control state detection stage and the overall system are broken down. The results for control state detection include the accuracy (ACC_1), positive predictive value (PPV), negative predictive value (NPV), true positive rate (TPR) and true negative rate (TNR) achieved by OSRD method during the online sessions. In this analysis, control state has been considered as the positive class and the non-control state as the negative class. Results of the overall system include the control state detection and command detection stages. Hence, control trials were considered correct whether the control state and the command were correctly determined at the same time. For the overall system, accuracy (ACC_2) and information transfer rate (ITR) are given. ITR measures the amount of information conveyed by a BCI system per unit of time [1]. ITR is valid in memoryless systems, where all possible selections are equally probable [24]. Our system meets both assumptions and thus,

TABLE II
OFFLINE TEST ACCURACIES

	No. Sequences							
	1		5		10		15	
	R	S	R	S	R	S	R	S
U01	55.0	54.2	84.2	85.0	98.3	98.3	99.2	99.2
U02	60.8	59.2	88.3	89.2	95.8	95.8	97.5	97.5
U03	63.3	64.2	94.2	95.0	97.5	97.5	100	100
U04	64.2	64.2	83.3	84.2	96.7	95.8	99.2	99.2
U05	61.7	60.8	87.5	87.5	96.7	96.7	97.5	98.3
U06	65.8	64.2	95.8	95.0	96.7	96.7	98.3	98.3
U07	56.7	58.3	75.0	75.8	79.2	76.7	87.5	86.7
U08	63.3	62.5	97.5	97.5	100	100	100	100
U09	61.7	60.0	89.2	88.3	91.7	92.5	95.0	97.5
U10	50.8	50.0	80.8	80.0	88.3	89.2	88.3	88.3
U11	65.0	65.8	91.7	91.7	100	100	100	100
U12	60.0	61.7	79.2	80.0	86.7	85.8	92.5	92.5
U13	65.0	65.0	79.2	78.3	90.0	90.0	95.0	95.0
U14	60.0	56.7	94.2	95.0	99.2	98.3	100	100
U15	73.3	72.5	96.7	96.7	99.2	99.2	100	100
Mean	**61.8**	**61.3**	**87.8**	**87.9**	**94.4**	**94.2**	**96.7**	**96.8**
±SD	5.0	5.2	7.0	6.9	5.8	6.2	4.1	4.2

Accuracies given in percentage (%), R: training and testing with real non-control observations, S: training with synthetic non-control observations but testing with real non-control observations.

TABLE III
ONLINE EXPERIMENT RESULTS

	N_s	Control State Detection					Overall	
		ACC_1	PPV	NPV	TPR	TNR	ACC_2	ITR
U01	10	100	100	100	100	100	91.7	12.4
U02	10	95.8	100	92.3	91.7	100	91.7	12.4
U03	8	100	100	100	100	100	95.8	16.8
U04	10	91.7	100	85.7	83.3	100	91.7	12.4
U05	12	91.7	100	85.7	83.3	100	83.3	8.7
U06	10	95.8	100	92.3	91.7	100	91.7	12.4
U07	15	95.8	92.3	100	100	91.7	95.8	9
U08	8	95.8	92.3	100	100	91.7	95.8	16.8
U09	12	95.8	92.3	100	100	91.7	95.8	11.2
U10	15	91.7	91.7	91.7	91.7	91.7	91.7	8.2
U11	10	95.8	100	92.3	91.7	100	95.8	13.4
U12	10	91.7	91.7	91.7	91.7	91.7	91.7	12.4
U13	8	100	100	100	100	100	100	18.5
U14	12	95.8	100	92.3	91.7	100	91.7	10.3
U15	10	95.8	92.3	100	100	91.7	83.3	10.5
Mean	**10.7**	**95.5**	**96.8**	**94.9**	**94.5**	**96.7**	**92.5**	**12.4**
±SD	2.1	2.8	3.9	5.2	5.8	4.1	4.4	2.9

N_s: number of sequences, Control State Detection: results of control state detection stage using OSRD method, ACC_1: accuracy (%) of the control state detection stage, PPV: positive predictive value (%), NPV: negative predictive value (%), TPR: true positive rate (%), TNR: true negative rate (%), Overall: results of the overall system including control state detection and command selection, ACC_2: accuracy of the overall system, ITR: information transfer rate (bits/min).

ITR is applicable. As in the offline sessions, the stimulation frequency was set to 5.71 Hz.

IV. DISCUSSION

In this study, we investigated the SSVEPs elicited by the RCP in order to detect the control state in ERP-based spellers. To characterize this phenomenon, two different analysis were performed. The first analysis assessed how the variation of the stimulation frequency affected the SSVEPs. As can be noticed in Fig. 3, lower stimulation rates trigger SSEVPs of higher power, being maximum at 2.67 Hz and considerably lower at 6.66 Hz and 8.00 Hz. In addition, harmonics are noticeable in the grand average for 2.10 Hz, 2.67 Hz and 3.54 Hz. Despite the fact that higher SSVEP powers could increase the accuracy of OSRD method, low stimulation rates would decrease the ITR of the system. Accordingly, a value of 5.71 Hz was chosen for the offline and online experiments. Regarding the spatial distribution of the SSVEPs, we can point out several insightful implications. SSVEPs reach higher amplitudes in the midline electrodes of the frontal, central and parietal areas of the brain (i.e., Fz, Cz, CPz, Pz) for lower fequencies (i.e., 2.10 Hz, 2.67 Hz, 3.64 Hz, 5.71 Hz). However, as the stimulation frequency increases, SSVEPs reach a higher power in the electrodes closer to the visual cortex, located in the occipital lobe (i.e., POz, PO7, PO8, Oz). These results are in accordance with previous studies [18], [25], and suggest that the spatial distribution of the electrodes to detect this waveform in ERP-based spellers should take into account the stimulation frequency. Regarding the second analysis, the presence of the SSVEP in overt mode (see Fig. 4) implies that the cause of this waveforms are the peripheral stimuli of the visual odd-ball paradigm. Covert mode only triggers transient ERPs, including the P300 potential. Moreover, the linear superposition of these two waveforms explains the shape of the EEG signal in control mode. The main implication of this finding is that SSVEPs and ERPs elicited by the RCP are originated by different mechanisms. Thus, OSRD method could be considered independent of the command detection stage, reducing the inter-session variability.

Regarding the offline experiments, no significant differences have been found between OSRD trained with real and synthetic non-control observations regardless the number of sequences (Wilcoxon signed-rank test, p-value > 0.05). Thus, synthetic non-control observations may be considered similar to real non-control observations. Furthermore, the use of synthetic non-control observations reduces the duration of calibration sessions by half because the acquisition of non-control trials is no longer needed. As could be expected, the greater the number of sequences considered, the higher the precision of the algorithm. It should be noted that 13 out of 15 participants reached an accuracy higher than 90%. Additionally, the proposed method only uses two features, which assures a similar performance with less training observations. These characteristics represent a great advantage in real-life BCI applications.

Table III presents the results of the online session. OSRD method achieved an average accuracy of 95.5% for control state detection. PPV, NPV, TPR and TNR values show that OSRD is more reliable when it comes to detecting the non-control state. This is useful in tasks where avoiding false positives is a critical issue. Additionally, Table III also includes the ITR and the accuracy of the overall system. These results show that participants were able to control the ERP-based speller asynchronously without compromising the overall system performance. In fact, 13 out of 15 participants reached an overall accuracy greater than 90%, achieving an average accuracy of 92.5%. Furthermore, this session demonstrated the ability of OSRD to work in real time.

Table IV shows a comparative between previous approaches of asynchronous ERP-based spellers and our system. It should be noted that it is difficult to make a direct comparison of these

TABLE IV
COMPARATIVE OF PREVIOUS ASYNCHRONOUS ERP-BASED SPELLERS.

Study	Paradigm	Method	Subjects	Performance
Zhang et al. 2008 [8]	Single cell	Probabilistic analysis of SVM scores	4 CS	ITR 15 bits/min FPR† 0.71 events/min
Aloise et al. 2011 [9]	RCP	ROC thresholding using LDA scores	11 CS	ITR 11.19 bits/min FPR† 0.26 events/min
Martínez-Cagigal et al. 2017 [10]	RCP	ROC thresholding using LDA scores	5 CS 16 MDS	ACC* 95.75% ACC* 84.14%
Tang et al. 2018 [26]	RCP	Thresholding using SWLDA scores	4 CS	ACC* 90.3%
Aydin et al. 2018 [6]	RBP	Thresholding using LDA output	10 CS	ITR 43.15 bits/min ACC* 93.27%
Martínez-Cagigal et al. 2019 [12]	RCP	ROC thresholding using LDA scores	5 CS 16 MDS	ACC* 92.30% ACC* 80.6%
Pinegger et al. 2016 [17]	RCP	Spectral features using SAM Hybrid features using HAM	21 CS	ACC† 79.5% ACC† 95.5%
Present study	**RCP**	**Detection of SSVEPs provoked by RCP**	**15 CS**	**ACC† 95.5%** **ACC* 92.5%** **ITR 12.4 bits/min** **FPR 0.14 events/min**

RCP: row-column paradigm, RBP: region based paradigm, SVM: support vector machine, LDA: linear discriminant analysis, SWLDA: step-wise LDA, SAM: spectral analysis method, HAM: hybrid analysis method, CS: control subjects, MDS: motor disabled subjects, FPR: false positive rate, ACC: accuracy, ITR: information transfer rate, *Overall system, †Control state detection

studies due to differences in stimulation paradigms, signal processing and experimental settings. As a first approach, Zhang et al. proposed a system based on a single cell paradigm that assessed the statistical differences between the output of a SVM fed with target and non-target epochs [8]. This study achieved an ITR $= 15$ bits/min and a FPR $= 0.71$ events/min with 4 control subjects (CS). Despite the suitable ITR, their FPR suggests that the asynchrony management could be improved. In fact, OSRD achieved a FPR $= 0.14$ events/min in the online session, showing the reliability of our approach. Aloise et al. improved the accuracy of the asynchronous detection, reaching a FPR $= 0.26$ events/min [9]. However, our approach still improves this value. Martínez-Cagigal et al. [10], [12] used the same asynchronous framework, obtaining with CS overall system accuracies of 95.75% and 92.30% respectively. Among the previous studies, Aydin et al. [6] reached the highest performance with an ITR $= 43.15$ bits/min. However, it should be taken into account that they used a different stimulation setting: the region based paradigm (RBP). This paradigm uses regions of commands far from each other and selections are performed in two levels to maximize speed, which makes it difficult to compare their results with the rest of the studies [6]. All previous studies rely on a threshold based on different parameters calculated with the output of the ERP classification stage to determine the control state (i.e., area under the ROC, probability analysis, etc). Despite their usefulness, these approaches present a number of limitations. Their performance is affected by high inter-session variability due to changes in the latency and amplitude of the ERPs [2], [5]. Moreover, in our own experience in [10], [12], a slight decrease in the ERP classification performance overrides the threshold, causing a drastic decrease in the detection accuracy of the control state. In addition, a high number of false negatives implies that correctly spelled trials could be lost. Independent methods of the ERP classification stage would help to overcome these limitations. As stated before, OSRD is included within these kind of approaches. Pinegger et al. [17] explored this strategy. In this study, SSVEPs elicited by the RCP were also used to determine the user's control state. They proposed the spectral analysis method (SAM), based on a threshold on the FFT values of averaged epochs of 1000 ms (256 samples). Despite the novelty of this method, its performance (TPR $= 88.3\%$, TNR $= 73.7\%$, ACC $= 79.5\%$, 15 sequences) suggests that the control state detection could be improved. SAM method is the only one in this comparative that allows a direct comparison with OSRD, since it relies on the same phenomenon to determine the control state. In this regard, OSRD method outperforms SAM in all the given metrics (TPR $= 94.5\%$, TNR $= 96.7\%$, ACC $= 95.5\%$, 10.7 sequences). In order to improve their results, Pinegger et al. combined SAM with the ERP classifier scores in the hybrid analysis method (HAM). This method achieved a similar performance to OSRD (TPR $= 99.0\%$, TNR $= 93.2\%$, ACC $= 95.5\%$, 15 sequences). However, HAM still depends on the ERP classification stage, having the same drawbacks that previous approaches. Unlike OSRD, all asynchrony methods included in this comparative need non-control trials to be calibrated, which doubles the duration of training sessions. Furthermore, given their dependence on the ERP classifier, they should be often recalibrated due to the high inter-session variability of the latency and shape of the ERPs [11]. Thus, this procedure would ultimately take a great amount of time from users, reducing the applicability of these methods in a real context. OSRD addresses this issue by creating synthetic non-control observations from control trials, increasing the usability of the system. Finally, it is worthy to mention that we did not include hybrid BCIs in the comparative, since the use of different control signals does not allow a direct comparison. In addition, hybrid BCIs require the user to manage with two or more control signals, increasing the complexity of the systems and compromising its usability. In fact, these BCIs are more demanding, and might be challenging for certain users

who cannot maintain high levels of concentration [6], [16].

Despite the positive results achieved in this study, we can point out several limitations. This study failed to test OSRD method with motor disabled people, who would be the target users. In this regard, BCI systems generally present lower performance with motor disabled subjects. Nevertheless, SSVEPs could be less affected than ERPs in this population [27]. OSRD calculates two features focused on detecting the SSVEPs provoked by the RCP. However, complementary metrics based on statistical analysis could improve the performance and should be explored in the future [22], [28]. Finally, in order to validate the efficacy of OSRD, a training set of 60 trials was used, even though smaller sets are more suitable in real-life applications.

V. Conclusion

This study presents a novel method to determine the user's control state by means of the detection of SSVEPs elicited by the RCP in ERP-based spellers. We demonstrated that these waveforms, whose shape and spatial distribution depend on the stimulation frequency, are provoked by peripheral stimuli of the RCP. The proposed method has been validated in offline and online experiments, achieving an average accuracy of 95.5% for control state detection and outperforming other related state of the art methods. Additionally, our approach presents two main advantages: it is independent of the ERP classification stage, reducing the inter-session variability; and it is the first asynchrony method that does not need to register non-control trials, drastically reducing the duration of calibration sessions. These features make OSRD a suitable method to implement asynchronous ERP-based spellers in a real context.

References

[1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain Computer Interfaces for communication and control," *Clinical neurophysiology*, vol. 4, no. 113, pp. 767–791, 2002.

[2] L. F. Nicolas-Alonso and J. Gomez-Gil, "Brain computer interfaces, a review," *Sensors*, vol. 12, no. 2, pp. 1211–1279, 2012.

[3] S. J. Luck, *An introduction to the event-related potential technique*. MIT press, 2014.

[4] L. A. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalography and Clinical Neurophysiology*, vol. 70, no. 6, pp. 510–523, 1988.

[5] F. Schettini, F. Aloise, P. Aricò, S. Salinari, D. Mattia, and F. Cincotti, "Self-calibration algorithm in an asynchronous P300-based brain-computer interface," *Journal of Neural Engineering*, vol. 11, no. 3, 2014.

[6] E. A. Aydin, O. F. Bay, and I. Guler, "P300-Based Asynchronous Brain Computer Interface for Environmental Control System," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 3, pp. 653 – 663, 2018.

[7] A. Rezeika, M. Benda, P. Stawicki, F. Gembler, A. Saboor, and I. Volosyak, "Brain–computer interface spellers: A review," *Brain Sciences*, vol. 8, no. 4, 2018.

[8] H. Zhang, C. Guan, and C. Wang, "Asynchronous P300-based brain-computer interfaces: a computational approach with statistical models." *IEEE transactions on bio-medical engineering*, vol. 55, no. 6, pp. 1754–63, 2008.

[9] F. Aloise, F. Schettini, P. Aricò, F. Leotta, S. Salinari, D. Mattia, F. Babiloni, and F. Cincotti, "P300-based brain-computer interface for environmental control: An asynchronous approach," *Journal of Neural Engineering*, vol. 8, no. 2, 2011.

[10] V. Martínez-Cagigal, J. Gomez-Pilar, D. Álvarez, and R. Hornero, "An Asynchronous P300-Based Brain-Computer Interface Web Browser for Severely Disabled People," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 8, pp. 1332–1342, 2017.

[11] S. He, R. Zhang, Q. Wang, Y. Chen, T. Yang, Z. Feng, Y. Zhang, M. Shao, and Y. Li, "A P300-Based Threshold-Free Brain Switch and Its Application in Wheelchair Control," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 6, pp. 715–725, 2017.

[12] V. Martínez-Cagigal, E. Santamaría-Vázquez, J. Gomez-Pilar, and R. Hornero, "Towards an accessible use of smartphone-based social networks through brain-computer interfaces," *Expert Systems with Applications*, vol. 120, pp. 155–166, 2019.

[13] R. C. Panicker, S. Puthusserypady, A. P. Pryana, and Y. Sun, "Asynchronous P300 BCI: SSVEP-based control state detection," *European Signal Processing Conference*, vol. 58, no. 6, pp. 934–938, 2010.

[14] Y. Li, J. Pan, F. Wang, and Z. Yu, "A hybrid BCI system combining P300 and SSVEP and its application to wheelchair control," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 11, pp. 3156–3166, 2013.

[15] Y. Yu, Z. Zhou, Y. Liu, J. Jiang, E. Yin, N. Zhang, Z. Wang, Y. Liu, X. Wu, and D. Hu, "Self-paced operation of a wheelchair based on a hybrid brain-computer interface combining motor imagery and P300 potential," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 12, pp. 2516–2526, 2017.

[16] S. Amiri, R. Fazel-Rezai, and V. Asadpour, "A Review of Hybrid Brain-Computer Interface Systems," *Advances in Human-Computer Interaction*, vol. 2013, pp. 1–8, 2013.

[17] A. Pinegger, J. Faller, S. Halder, S. C. Wriessnegger, and G. R. Müller-Putz, "Control or non-control state: that is the question! An asynchronous visual P300-based BCI approach," *Journal of Neural Engineering*, vol. 12, no. 1, p. 014001, 2015.

[18] F. B. Vialatte, M. Maurice, J. Dauwels, and A. Cichocki, "Steady-state visually evoked potentials: Focus on essential paradigms and future perspectives," *Progress in Neurobiology*, vol. 90, no. 4, pp. 418–438, 2010.

[19] E. Santamaría-Vázquez, V. Martínez-Cagigal, and R. Hornero, "MEDUSA: Una nueva herramienta para el desarrollo de sistemas Brain-Computer Interface basada en Python," *Cognitive Area Networks*, vol. 5, no. 1, pp. 87–92, 2018.

[20] W. Krzanowski, *Principles of multivariate analysis*. OUP Oxford, 2000.

[21] Z. Lin, C. Zhang, W. Wu, and X. Gao, "Frequency Recognition Based on Canonical Correlation Analysis for SSVEP-Based BCIs," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 12, pp. 2610–2614, 2006.

[22] G. Bin, X. Gao, Z. Yan, B. Hong, and S. Gao, "An online multi-channel SSVEP-based brain-computer interface using a canonical correlation analysis method," *Journal of Neural Engineering*, vol. 6, no. 4, 2009.

[23] K. Fukunaga, *Introduction to statistical pattern recognition*. Academic press, 2013.

[24] W. Speier, C. Arnold, and N. Pouratian, "Evaluating True BCI Communication Rate through Mutual Information and Language Models," *PLoS ONE*, vol. 8, no. 10, 2013.

[25] R. Srinivasan, F. A. Bibi, and P. L. Nunez, "Steady-state visual evoked potentials: Distributed local sources and wave-like dynamics are sensitive to flicker frequency," *Brain Topography*, vol. 18, no. 3, pp. 167–187, 2006.

[26] J. Tang, Y. Liu, J. Jiang, Y. Yu, D. Hu, and Z. Zhou, "Toward Brain-Actuated Mobile Platform," *International Journal of Human-Computer Interaction*, vol. 00, no. 00, pp. 1–12, 2018.

[27] A. Combaz, C. Chatelle, A. Robben, G. Vanhoof, A. Goeleven, V. Thijs, M. M. Van Hulle, and S. Laureys, "A Comparison of Two Spelling Brain-Computer Interfaces Based on Visual P3 and SSVEP in Locked-In Syndrome," *PLoS ONE*, vol. 8, no. 9, 2013.

[28] V. Martínez-Cagigal, E. Santamaría-Vázquez, and R. Hornero, "Asynchronous Control of P300-Based Brain–Computer Interfaces Using Sample Entropy," *Entropy*, vol. 21, no. 3, p. 230, 2019.