

# SCIENTIFIC REPORTS



OPEN

## Code-modulated visual evoked potentials using fast stimulus presentation and spatiotemporal beamformer decoding

Benjamin Wittevrongel<sup>1</sup> , Elia Van Wolputte<sup>2</sup>  & Marc M. Van Hulle<sup>1</sup>

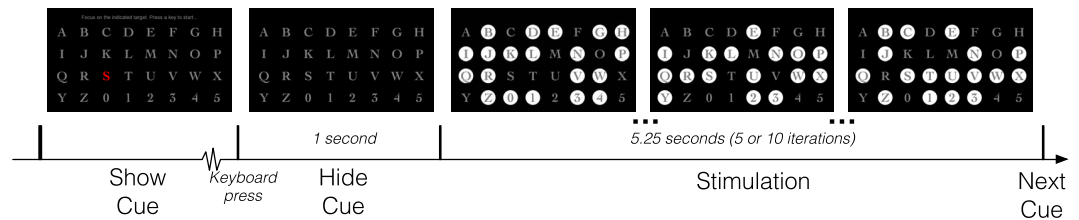
When encoding visual targets using various lagged versions of a pseudorandom binary sequence of luminance changes, the EEG signal recorded over the viewer's occipital pole exhibits so-called code-modulated visual evoked potentials (cVEPs), the phase lags of which can be tied to these targets. The cVEP paradigm has enjoyed interest in the brain-computer interfacing (BCI) community for the reported high information transfer rates (ITR, in bits/min). In this study, we introduce a novel decoding algorithm based on spatiotemporal beamforming, and show that this algorithm is able to accurately identify the gazed target. Especially for a small number of repetitions of the coding sequence, our beamforming approach significantly outperforms an optimised support vector machine (SVM)-based classifier, which is considered state-of-the-art in cVEP-based BCI. In addition to the traditional 60 Hz stimulus presentation rate for the coding sequence, we also explore the 120 Hz rate, and show that the latter enables faster communication, with a maximal median ITR of 172.87 bits/min. Finally, we also report on a transition effect in the EEG signal following the onset of the stimulus sequence, and recommend to exclude the first 150 ms of the trials from decoding when relying on a single presentation of the stimulus sequence.

Among the gamut of brain-computer interfacing (BCI) paradigms<sup>1</sup>, the code-modulated visual evoked potential (cVEP) has been reported to yield one of the highest information transfer rates (ITRs)<sup>2</sup>. The cVEP paradigm defines a binary sequence of high and low stimulus intensities with unequal duty cycles, called the 'code', and uses for each selectable target a unique lagged version of this code<sup>2-4</sup>. As coding sequence, the m-sequence<sup>5</sup> is often chosen because of its favourable autocorrelation properties<sup>2</sup>, amongst other properties<sup>6</sup>. M-sequences have also been applied in other fields such as fMRI<sup>7</sup> and sensor technology<sup>8</sup>. Albeit rarely adopted, as an alternative to m-sequences, a periodic pseudorandom binary code has been described for cVEP<sup>9</sup>.

As far as we are aware, only one group previously investigated the application of a faster stimulus presentation for cVEP (i.e., higher than the traditional 60 Hz)<sup>10-13</sup>. The encoding sequence was presented to the subject by LEDs at a carrier frequency of 40 Hz controlled by an Arduino micro-controller, which would compare to a 80 Hz screen refresh rate. The authors investigated the effect of stimulation colour<sup>10,13</sup>, classifier kernels<sup>11</sup> and filter bands<sup>12</sup>, but could not achieve a higher decoding performance for the faster stimulus rate. In one of their studies<sup>11</sup>, they report on the decoding performance with an increasing number of m-sequence repetitions, but did not consider the implication on the performance in terms of ITR.

Although cVEP has achieved among the highest ITR, the paradigm is considerably less studied compared to other visual BCI paradigms such as the P300 event-related potential (ERP) and the steady-state visual evoked potential (SSVEP) (see the review of Gao and coworkers<sup>14</sup> for reference). Traditionally, cVEPs are decoded from electroencephalography (EEG) using a template matching algorithm<sup>2</sup>, canonical correlation analysis (CCA)<sup>13</sup> or a combination<sup>15,16</sup>. BCIs adopting these algorithms have been proven successful in online settings, including EEG-based spelling applications<sup>17</sup> and robot control<sup>18,19</sup>, and have also been applied in an intracranial EEG (iEEG) setting<sup>20</sup>. In recent research, the support vector machine (SVM) has been shown to identify targets more accurately than the traditional decoding algorithms<sup>10</sup>, with a linear kernel achieving in the highest accuracy<sup>11</sup>.

<sup>1</sup>Department of Neurosciences, KU Leuven, Leuven, Belgium. <sup>2</sup>Department of Computer Science, KU Leuven, Leuven, Belgium. Correspondence and requests for materials should be addressed to B.W. (email: [benjamin.wittevrongel@kuleuven.be](mailto:benjamin.wittevrongel@kuleuven.be))



**Figure 1.** Time-course of one trial during the experiment.

Session	Stimulation			Analysis	
	stimulus presentation rate	m-sequence repetitions per trial	trial duration	downsampling	samples per segment
$S_{60}$	60 Hz	5	5.25 s	100 Hz	105
$S_{120}$	120 Hz	10	5.25 s	200 Hz	105

**Table 1.** Stimulation and analysis details of both sessions. For the faster stimulation rate, the downsampling rate is doubled, leading to an equal number of samples in the segments of both sessions. Note that a segment corresponds to the EEG response elicited by one full presentation of the m-sequence.

Recently, a spatiotemporal extension of the beamforming algorithm has been introduced in EEG-BCI and shown to yield promising results with EEG signals that have consistent spatial and temporal characteristics, such as the N400<sup>21</sup> and P300<sup>22</sup> ERPs. With SSVEP, the combination of the spatiotemporal beamformer with a time-domain analysis<sup>23,24</sup> proved successful in both offline<sup>25</sup> and online<sup>26</sup> settings.

The goal of this study is to assess the performance of the spatiotemporal beamforming algorithm for target identification when using cVEP-based encoding, and to compare the performance for both traditional (60 Hz) and high-speed (120 Hz) stimulus presentations.

## Methods

**Subjects.** Seventeen subjects with normal or corrected-to-normal vision participated in the experiment (14 female, 13 right handed, aged  $22.35 \pm 2.9$ , ranging from 18 to 30 years old). Prior to the experiment, the subjects read and, when they agreed, signed an informed consent form approved by the ethical committee of our university hospital UZ Leuven. All subjects received a monetary reward for their participation. This study was carried out in accordance with the relevant guidelines and regulations.

**Experimental design.** The interface consisted of 32 circular white targets (4 cm diameter, 2 cm vertical and horizontal inter-target distance) that follow an m-sequence stimulation paradigm (see further) and that were overlaid with static (i.e., non-flickering) grey letters or numbers arranged in a matrix (Fig. 1). The interface was presented on a ViewPixx-EEG monitor (24 inch, native 120 Hz refresh rate, resolution of  $1920 \times 1080$ , VPixx Technologies, Canada). The subjects were seated approximately 60 cm from the monitor. At this distance, the circular targets spanned a visual angle of approximately  $3.8^\circ$ , with an inter-target angle of  $1.9^\circ$ . The experiment was implemented in Matlab, using the Psychophysics Toolbox extensions<sup>27–29</sup>.

The following m-sequence of length of 63 was used to encode the targets:

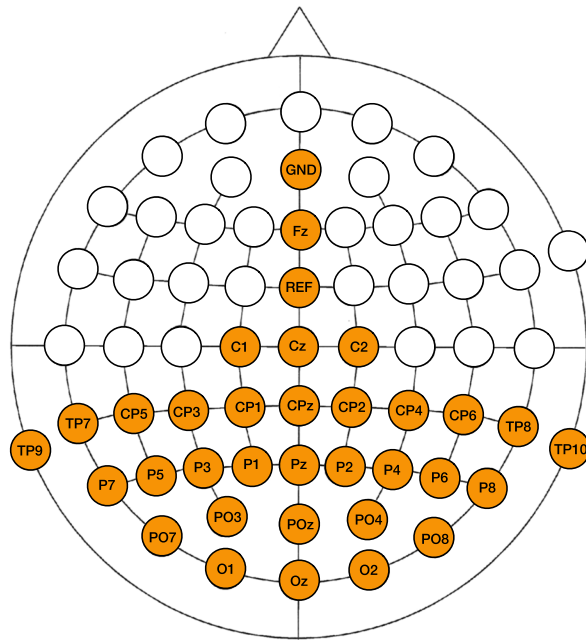
000100001011001010100100111100.00011011100110001110101111101101

where targets were lagged by integer multiples of two frames. We adopted the equivalent-neighbours strategy used in other studies<sup>15,17</sup>, but decided not to implement the additional outer border in order to reduce visual demand.

Figure 1 visualises the experimental interface during one trial. A trial started with the presentation of a cue (i.e., one target shown in red). Subjects were asked to redirect their gaze to the cued target and to press a button to start the stimulation. After that, all targets were hidden (with the characters still shown in grey) for one second, followed by the stimulation phase during which all targets adopted their unique lagged m-sequence and repeated this sequence either 5 or 10 times (depending on the session, see further). To avoid visual fatigue and boredom, subjects were allowed to take breaks between trials.

Unlike traditional 60 Hz monitors, the monitor used in our experiment had a refresh rate of 120 Hz, which allowed us to experiment with high-speed presentations of the coding sequence. The full experiment consisted of two sessions. In one session,  $S_{120}$ , the stimulus presentation followed the screen refresh rate. In the other session,  $S_{60}$ , we simulated the stimulation as it would be presented on a 60 Hz screen by presenting each entry of the m-sequence for two frames before moving to the next entry. In each trial of  $S_{120}$  and  $S_{60}$ , the m-sequence was repeated 10 and 5 times, respectively. In both sessions, the stimulation duration per trial was 5.25 seconds, and all targets were cued 5 times in pseudorandom order, leading to a total of 160 ( $=32 \times 5$ ) trials per session.

The two sessions were counterbalanced across subjects: 9 of the 17 subjects started with  $S_{60}$ , while the other 8 performed  $S_{120}$  first. Table 1 summarises the details of the two sessions.



**Figure 2.** Locations of the 32 electrodes used during the experiment.

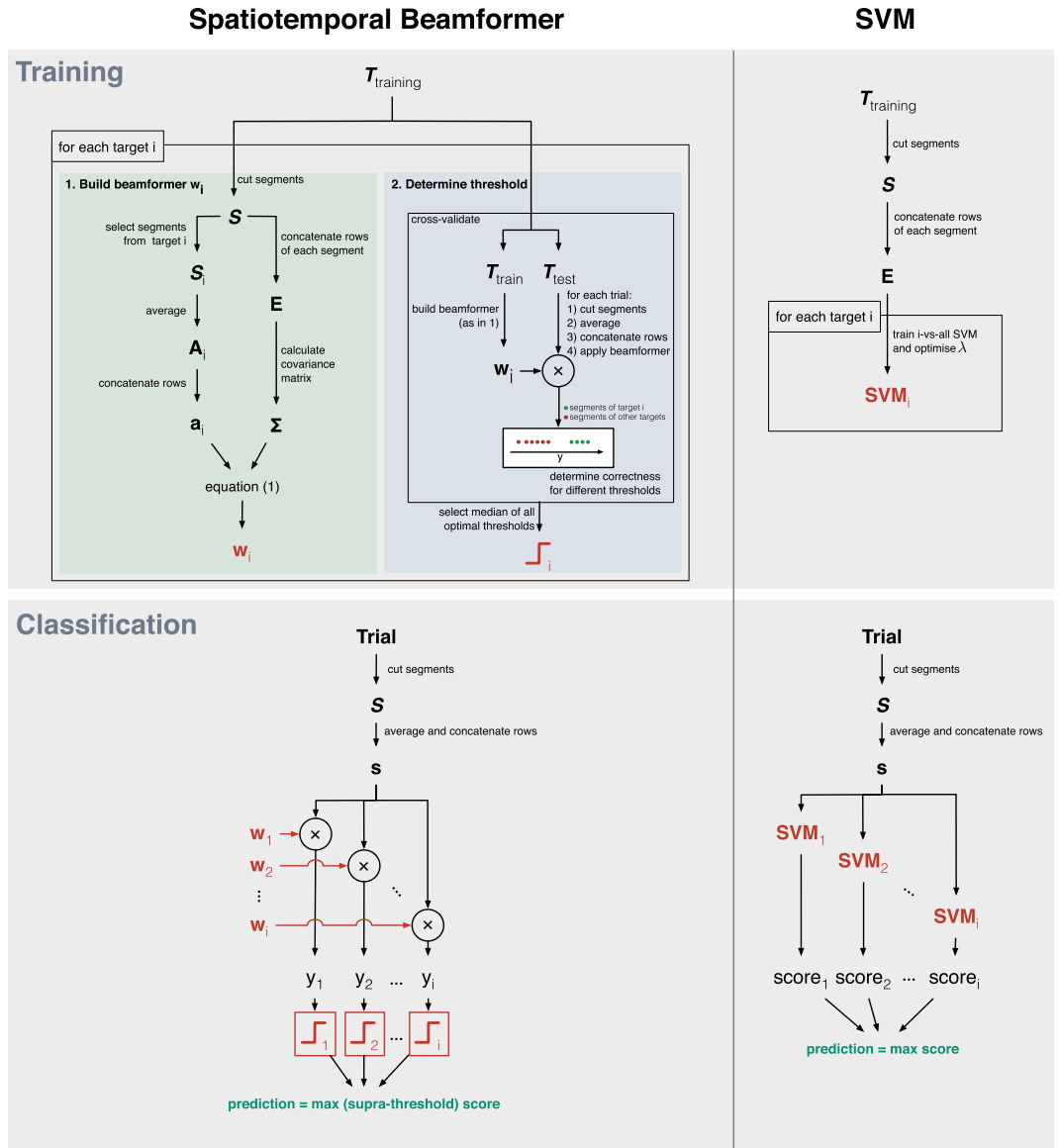
**Recording.** EEG was recorded continuously using a SynampsRT device (Compumedics Neuroscan, Australia) operating at a sampling rate of 2000 Hz with 32 active Ag/AgCl electrodes covering the parietal and occipital poles, where consistent activations in response to cVEP stimulation are expected<sup>10,15</sup>. The ground (GND) and reference (REF) electrodes were located at AFz and FCz, respectively (Fig. 2). Conductive gel was applied at each electrode site and impedances were kept below 2 k $\Omega$ .

**Preprocessing.** The raw signal was re-referenced offline to the average of both mastoids signals (TP9 and TP10) and filtered between 4 and 31 Hz using a 4th order Butterworth filter, in order to attenuate the presence of artefacts such as slow drifts due to electrode gel expiration and sweat, low frequency oscillations due to electrode movements, high-frequency extraphysiologic noise, and powerline interference. The EEG was then cut into 5.25-second epochs starting from the onset of the stimulation, and labeled with the corresponding cued target. Finally, the epochs of  $S_{60}$  and  $S_{120}$  were downsampled to 100 Hz and 200 Hz, respectively, and stored for further analysis. The difference in downsampling rate was included to obtain a fair comparison between the classification results (each repetition of the m-sequence at both the traditional and faster stimulus rate has an equal number of samples, see Table 1). For each subject and session, 160 labeled epochs were extracted and saved.

**Classification.** Target identification was achieved using a classifier based on the linearly-constrained minimum-variance (LCMV) spatiotemporal beamformer<sup>21</sup>. This recent extension of the original spatial beamformer estimates the contribution of an a-priori specified activation pattern (i.e., a template, a signal of interest) to the current input. It has been shown that LCMV beamforming is a special case of Minimum Variance Distortionless Response (MVD) beamforming<sup>30</sup>, introduced to improve the robustness of the latter<sup>31</sup>. The EEG responses to the stimuli of interest are not only confluent by ongoing brain activity but can also be modulated by the subject's attention level, motivation and fatigue. The LCMV beamformer in an EEG context has shown to be effective as spatial filter for ERP detection<sup>32</sup> and source localisation for studying source connectivity<sup>33–35</sup>, and its spatiotemporal extension has shown effective for ERP analysis<sup>21</sup> and as target identification algorithm in BCI settings<sup>22,25,26</sup>.

Since each target elicits a different brain response (cf. unique lags of the m-sequence), each target evokes a unique EEG activation pattern, and training the classifier thus involves the estimation of 32 activation patterns, each used to construct a beamformer tailored to a specific target. The training and classification procedures for both the beamformer- and SVM-based classifiers are depicted in Fig. 3.

**Beamforming.** The activation patterns and the beamformers (one for each target) were calculated from the training data  $T_{training} \in \mathbb{R}^{m \times t \times l}$ , where  $m$  is the number of channels,  $t$  is the number of samples and  $l$  is the number of epochs, as follows. For each epoch in  $T_{training}$ , a maximal number of  $c$ -second consecutive non-overlapping segments were extracted, where  $c$  represents the time needed to display one complete m-sequence. Let  $S \in \mathbb{R}^{m \times n \times r}$  be all  $r$  segments extracted and  $S_i \in \mathbb{R}^{m \times n \times k}$  be the segments from  $S$  in response to the cued target  $i \in [1..32]$ , with  $n$  the number of samples per segment and  $k$  the total number of segments extracted for target  $i$ . Note that, while the m-sequences of  $S_{120}$  only span half the time ( $c = 0.525$  s) compared to  $S_{60}$  ( $c = 1.05$  s), their sampling rate is doubled so that the segments obtained from epochs of both  $S_{60}$  and  $S_{120}$  have the same number of samples.



**Figure 3.** Visual representation of the training and classification procedure for the beamformer- and SVM-based classifier.

The spatiotemporal activation pattern  $A_i \in \mathbb{R}^{m \times n}$  for target  $i$  was then obtained as the average of all  $k$  segments from  $S_i$ . The spatiotemporal beamformer  $w_i \in \mathbb{R}^{(mn) \times 1}$  for target  $i$  was calculated as an LCMV beamformer as follows: let  $E \in \mathbb{R}^{r \times (mn)}$  be the matrix where each row is obtained by concatenating the rows of a corresponding sequence  $S[\ast, \ast, r]$ ,  $\Sigma \in \mathbb{R}^{(mn) \times (mn)}$  the covariance matrix of  $E$ , and  $a_i^T \in \mathbb{R}^{1 \times (mn)}$  a vector containing the concatenated rows of  $A_i$ . The LCMV beamformer under constraint  $a_i^T w_i = 1$  can be calculated using the method of Lagrange multipliers<sup>36</sup>:

$$w_i = \frac{\Sigma^{-1} a_i}{a_i^T \Sigma^{-1} a_i} \quad (1)$$

and applied to the data as a simple weighted sum:  $y_i = s w_i$ , where  $s \in \mathbb{R}^{1 \times (mn)}$  indicates the concatenated rows of an input segment  $S_{in} \in \mathbb{R}^{m \times n}$ .

In our study, the covariance matrix  $\Sigma$  was estimated using Matlab's (2015a) *cov* function and was inverted using the *pinv* function, which calculates the Moore-Penrose pseudoinverse, to account for possible singularity of  $\Sigma$ .

In some studies, a single activation pattern  $A_1$  was calculated based on the EEG response to target 1, and the activation pattern  $A_i$  of target  $i$  was constructed as a circular-shifted version of  $A_1$  (following the phase difference between the  $m$ -sequences of targets 1 and  $i$ )<sup>2,15</sup>. However, given the availability of training data for each target, we opted to calculate the activation patterns for each target independently. In this way, discontinuities introduced by the circular shift were avoided and minor variations between templates were taken into account, leading to more accurate beamformers.

**Classifier.** In addition to building a beamformer for each target, a threshold was determined for each target in order to classify segments (in a one-vs-all fashion) into target-(positive class) and non-target (negative class). The threshold for each target was optimised via a Receiver Operating Characteristic (ROC) analysis<sup>37,38</sup>, using an additional 4-fold cross-validation on the training data (3 folds were used to train the beamformer, the remaining fold to test its performance). The ROC curve plots binary classification performance as a function of threshold value. Since the maximum classification performance could be reached for multiple thresholds (equal ROC points or points on the maximal iso-performance line), we selected the median of these.

Classification of a new epoch involved the extraction of the segments  $S_{test}$ , using an identical procedure as for the training epochs. The segments were averaged and concatenated and then independently filtered by each beamformer to obtain a score  $y_i$  for each target  $i$ . Among the scores that exceeded the corresponding threshold, the one with the highest score was taken as winner. In case of none of the scores exceeded their threshold, the winner was determined by the highest (sub-threshold) score.

We compared our classifier based on spatiotemporal beamforming (stBF) with a SVM-based classifier. Similar to before, segments are extracted from the training epochs and concatenated to form feature vectors (cfr. *E*). Then, for each target, a one-vs-all linear SVM<sup>39</sup> was trained, whose regularisation parameter  $\lambda$  was optimised using a line-search strategy and 4-fold cross validation<sup>40</sup>. All SVMs were trained using the modified finite Newton method<sup>41</sup>. This procedure was successfully applied before to detect the P300 event-related potential (ERP) in patients with incomplete locked-in syndrome (LIS)<sup>42</sup>, to detect error-related potentials (ErrPs) in healthy subjects<sup>43</sup>, and served as a comparison for the spatiotemporal beamformer for P300 detection<sup>22</sup>. SVMs have been shown to outperform the traditional CCA classifier for cVEP detection<sup>10,11</sup>, and we opted for an optimised version of the SVMs in order to maximise accuracy. Prediction of a given (concatenated) input segment (cfr. *s*) was given by the SVM returning the highest (i.e., most positive) score.

**Channel selection.** For each subject, the channels included in the analysis were obtained using a greedy approach, in which we iteratively added the channel that improved the accuracy the most until it did no longer improve or until 100% accuracy was reached. As optimisation criterion, we used the classification accuracy obtained with the beamformer-based classifier when averaging two repetitions of the m-sequence (i.e., signal length of 2.10 and 1.05 sec for  $S_{60}$  and  $S_{120}$ , respectively).

**Transition effect.** It has been shown that the brain exhibits a latency of 100 to 150 ms in response to SSVEP stimulation<sup>44,45</sup>. During this time, the SSVEP is not stable, and in SSVEP-BCI research, the initial 100 to 150 ms of the epochs (time-locked to the onset of the flickering stimulation) is often excluded from analysis as it leads to increased accuracies<sup>25,46</sup>. Similar to SSVEP, cVEP is a visual paradigm adopting flickering stimulation (albeit not periodic), and we tested whether performance could be improved by excluding the initial 150 ms of each epoch. Note that, when excluding the first 150 ms of each epoch, an additional 150 ms is required at the end of the epoch to obtain the same number of complete m-sequences. For example, when excluding the initial signal, the first full m-sequence requires  $0.150 + 1.05 = 1.2$  seconds, compared to just 1.05 second without the exclusion.

In this study, we ran the analysis both with and without the exclusion of the initial 150 ms of each epoch.

**Performance evaluation.** The performance of the classifiers was estimated offline using a stratified 5-fold cross-validation strategy. Since each target was cued 5 times, each fold contained one 5.25-second epoch for each target. We obtained the target identification accuracy for different signal lengths, corresponding to multiples of the time needed to present one repetition of the m-sequence.

As the two stimulus presentation rates as well as the possible exclusion of the initial signal lead to differences in stimulation time, one should be careful in interpreting the accuracies obtained by the different conditions. The ITR, however, takes into account the stimulation length and therefore provides a fair comparison between the conditions. Hence, next to target identification accuracy, we also measure ITR (in bits/min) as follows<sup>47,48</sup>:

$$ITR = \frac{\log_2 N + p \cdot \log_2 p + (1 - p) \cdot \log_2 \left( \frac{1-p}{N-1} \right)}{t/60}, \quad (2)$$

where  $N$  is the number of selectable targets,  $p$  is the accuracy of target identification, and  $t$  is the time needed to make a selection (in seconds). In our study,  $N$  was equal to 32 and  $t$  was set to the stimulation length plus an additional 500 ms to account for the time the subject would need to switch their gaze to the next target. In the literature, studies investigating BCI spelling interfaces often adopt a gaze-switching interval in the range from 300 to 1000 ms<sup>17,46,49–51</sup>, and a 500 ms interval has been shown feasible in an online setting<sup>52</sup>, albeit with the SSVEP paradigm.

In addition to accuracy and ITR, we also measured the time needed to train the spatiotemporal beamformer- and SVM-based classifiers on all data for each subject. Timings were collected on a quad-core 2.3 GHz Intel i7 machine.

**Statistics.** Since the distributions do not consistently follow a gaussian distribution, we adopted the non-parametric (two-tailed) Wilcoxon signed rank test. We used this test to compare the accuracy of both classifiers and to compare the influence of excluding the first 150 ms of each epoch. The significance threshold was set to 0.05.

**Data Availability.** The (anonymised and pre-processed) data that support the findings of this study, as well as the implementation of the classifiers and the analysis, are made available at <https://kuleuven.box.com/v/CVEP>.

## Results

All results of  $S_{60}$  and  $S_{120}$  are summarised in Figs 4 and 5, respectively.

For both sessions, the optimal channel set was obtained using a greedy approach. For  $S_{60}$ , all subjects reached convergence with three or less channels (Fig. 4b), while between 3 and 6 channels were selected for  $S_{120}$  (Fig. 5b) before convergence was reached. The occipital channels Oz and O1 were selected most often (Figs 4a and 5a), and several parietal channels were selected by a smaller number of subjects, indicating considerable inter-subject variability.

Using the individually optimised channel sets, the target identification accuracy for both the spatiotemporal beamformer- and the SVM-based classifier are shown in Fig. 4c for  $S_{60}$  and Fig. 5c for  $S_{120}$ , both with and without the exclusion of the initial 150 ms of the stimulation. As expected, longer stimulation times (i.e., more repetitions of the m-sequence) increases performance. For the same stimulation lengths, the faster stimulus presentation ( $S_{120}$ ) is able to present twice the amount of m-sequences compared to  $S_{60}$ , which results in a higher accuracy for equal-length stimulation. Only the faster stimulus presentation in combination with the exclusion of the initial signal is able to surpass the accuracy threshold of 70% with a single repetition of the m-sequence. All other conditions require at least two repetitions to reach this threshold, which is deemed minimal for establishing reliable communication<sup>42,53–55</sup>.

Using the full signal, the accuracies of both classifiers differ significantly when averaging up to four ( $S_{60}$ , Fig. 4c) ( $p = 0.033$ ;  $p = 0.012$ ;  $p = 0.0017$  and  $p = 0.030$ ) and two ( $S_{120}$ , Fig. 5c) ( $p = 0.016$  and  $p = 0.020$ ) repetitions of the m-sequence, respectively. With the exclusion of the initial 150 ms, the two classifiers are not significantly different. However, within stBF, the accuracies with and without the exclusion of the initial signal significantly differ for 2 to 4 repetitions ( $S_{60}$ ) ( $p < 0.001$ ,  $p = 0.003$  and  $p = 0.004$ ) and 1 repetition ( $S_{120}$ ) ( $p < 0.001$ ) of the m-sequence, respectively. Similarly, within the SVM-based classifier, the accuracies with and without the exclusion of the initial signal significantly differ for one repetition ( $S_{60}$ ) ( $p < 0.001$ ) and 1 to 3 repetitions ( $S_{120}$ ) ( $p < 0.001$ ,  $p < 0.001$  and  $p = 0.008$ ) of the m-sequence, respectively.

A detailed inspection of the accuracy increase with one repetition of the m-sequence (Fig. 4d for  $S_{60}$  and Fig. 5d for  $S_{120}$ ) shows a negative relation between the increase in accuracy and the number of selected channels. This effect is most prominent for stBF at the traditional 60 Hz stimulus rate (Fig. 4d). All subjects requiring three channels have a reduction in accuracy by removing the first 150 ms of the epochs, while the other subjects have an increased accuracy. The SVM is less influenced by the number of channels, and removing the initial 150 ms signal only decreases its accuracy for two subjects. While this negative trend can also be detected for the faster stimulation rate, all subjects have an increased accuracy compared to when the initial 150 ms signal is included in the analysis.

For both sessions, the time needed to train stBF on all data of each subject is significantly lower compared to SVM (Fig. 4e for  $S_{60}$  and Fig. 5e for  $S_{120}$ ), and for both classifiers, the training time increases when more channels are included in the analysis.

For both the traditional and faster stimulus presentation rates, the median ITR reaches its maximal value of 100.46 and 172.87 bits/min, respectively, using the beamformer-based classifier, two repetitions of the m-sequence and the full signal (stimulation time = 2.1 and 1.05 seconds, respectively).

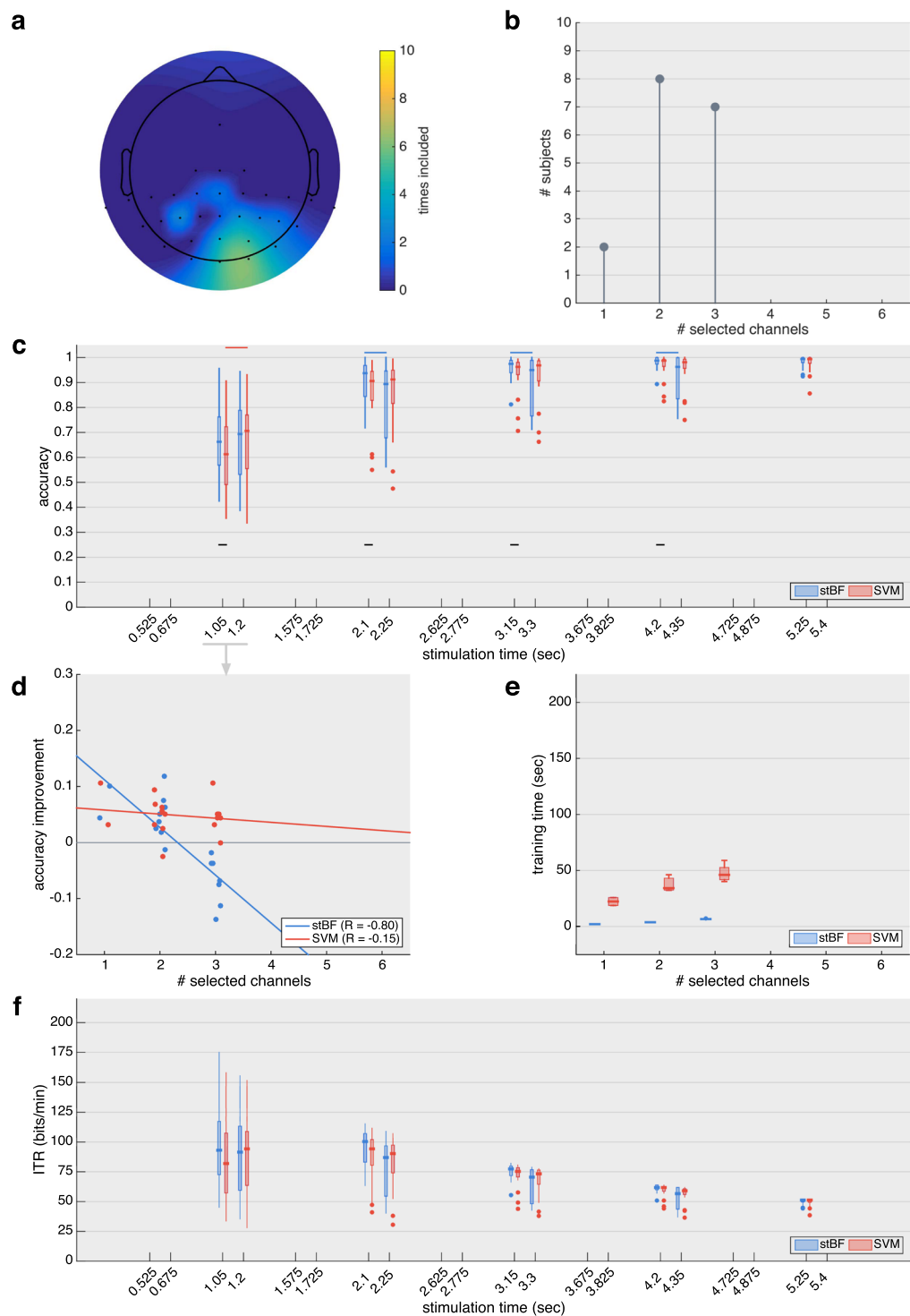
## Discussion

In this study, we assessed the feasibility of spatiotemporal beamforming for resolving m-sequence encoded targets in a cVEP setting, and investigated the influence of stimulus presentation rate on target identification accuracy and ITR.

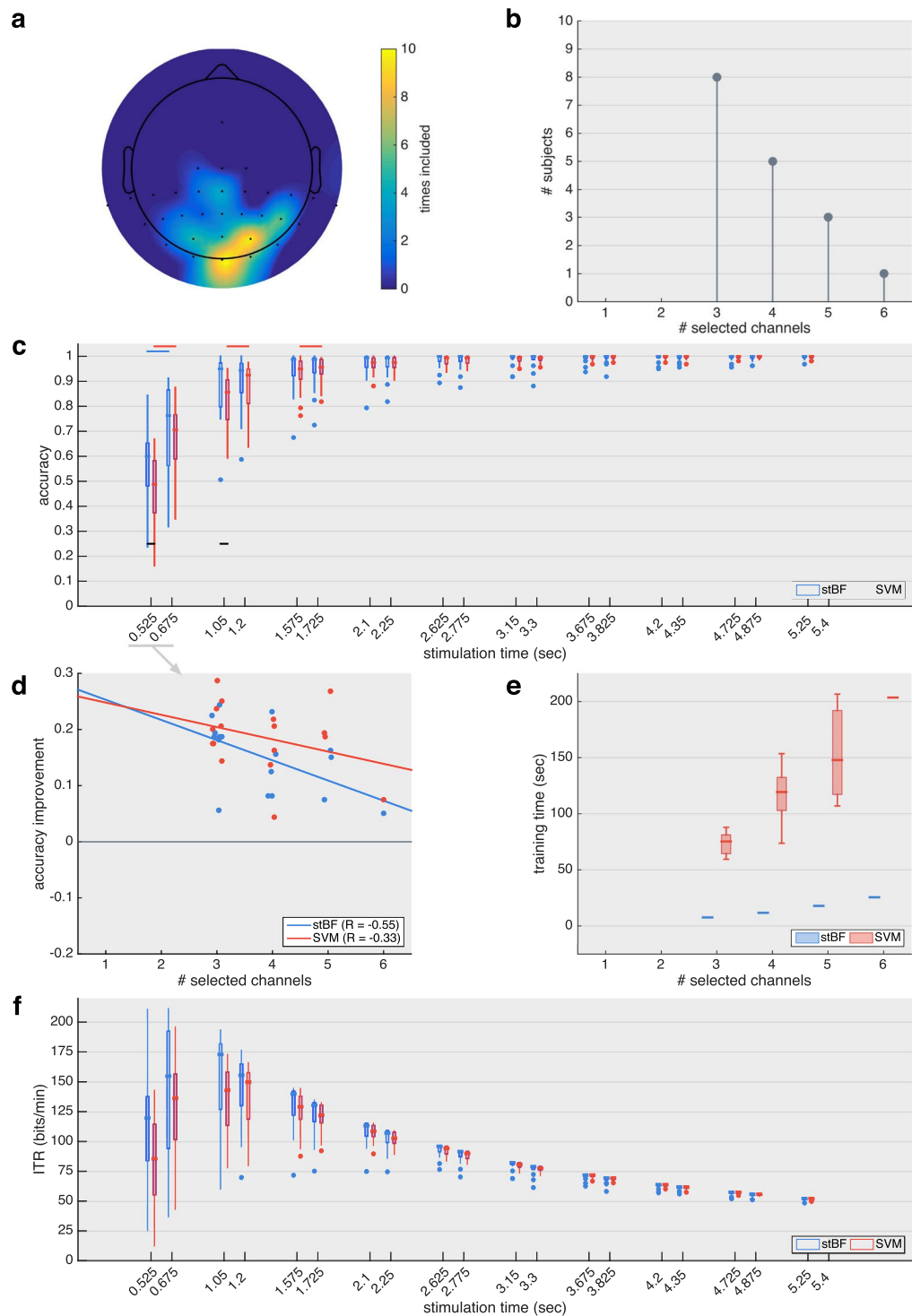
We showed that the proposed classifier is able to accurately discriminate targets, and that it is able to compete with a classifier based on optimised linear SVMs. We additionally show that a faster stimulus presentation rate is beneficial for the communication speed, as more iterations of the m-sequence can be presented in an equal amount of time. Both stimulation rates have similar performance in terms of number of m-sequence repetitions, and at least two repetitions are necessary to obtain a performance over 70%, which is deemed minimal for establishing reliable communication<sup>42,53–55</sup>. With two repetitions of the m-sequence, the median ITR is maximal and reaches 100.46 bits/min for the traditional 60 Hz and 172.87 bits/min for the faster 120 Hz stimulus presentation, respectively. As far as we are aware, no other cVEP study has reported a higher ITR. As commercial monitors with high frame rates are becoming increasingly more accessible at affordable prices, they are recommended for cVEP-BCI applications.

Compared to the SVM, the spatiotemporal beamformer can be trained considerably faster, as there are no parameters to be optimised. This could be important to achieve fast, online retraining of the beamformer-based classifier without causing the interface to be temporarily unavailable or to interfere with stimulation. The shorter training time would also allow for other optimisation algorithms to be executed (eg., channel selection, down-sampling rate, filtering range, etc.) that would otherwise not be able to complete within a reasonable time.

We present evidence that the cVEP response exhibits a transition effect following the onset of a stimulation sequence. Previously, a response latency of 100 to 150 ms has been described for SSVEP<sup>45</sup>, and in recent SSVEP-BCI research, the initial signal was excluded from the analysis to improve target identification<sup>25,46</sup>. In this study, excluding the initial 150 ms of each epoch improves classification accuracy of both classifiers when using merely one repetition of the m-sequence. The performance increase is negatively correlated with the number of selected channels and mostly affects the spatiotemporal beamformer, even causing a performance decrease when adopting three channels at the traditional 60 Hz stimulus presentation rate. For the 120 Hz case, all accuracies increase despite larger channels sets. The discrepancy between these results could be due to the fact that, when excluding the initial signal from each epoch, the last m-sequence of each epoch is not complete, and the number of complete training segments is reduced by 20% for  $S_{60}$  compared to only 10% for  $S_{120}$ . In order to maintain the same number of training segments, one could extend the stimulation of the training session by



**Figure 4.** Results for the session adopting the traditional 60 Hz stimulus rate. **(a)** Scalp plot indicating how many times each channel was selected by the greedy channel-selection algorithm across subjects. Note that most of the frontal and temporal area was not recorded during the experiment. **(b)** Summary of the total number of channels selected by the greedy channel-selection algorithm before convergence. **(c)** Target identification accuracy for both classifiers with an increasing number of repetitions of the stimulation sequence (1 m-sequence = 1.05 sec), with and without the initial 150 ms of each epoch. Black horizontal lines indicate significant differences between the classifiers. Blue and red horizontal lines indicate significant differences when excluding the first 150 ms. **(d)** Regression analysis of the increase in target identification accuracy based on one repetition of the m-sequence when excluding the first 150 ms. **(e)** Time needed to train the classifiers on all data of each subject. **(f)** Virtual ITR achieved when factoring in 0.5 seconds for gaze shifting.



**Figure 5.** Results for the session adopting the 120 Hz stimulus rate. **(a)** Scalp plot indicating how many times each channel was selected by the greedy channel-selection algorithm across subjects. Note that most of the frontal and temporal area was not recorded during the experiment. **(b)** Summary of the total number of channels selected by the greedy channel-selection algorithm before convergence. **(c)** Target identification accuracy for both classifiers with an increasing number of repetitions of the stimulation m-sequence (1 sequence = 0.525 sec), with and without the exclusion of the initial 150 ms of each epoch. Black horizontal lines indicate significant differences between the classifiers. Blue and red horizontal lines indicate significant differences when excluding the first 150 ms. **(d)** Regression analysis of the increase in target identification accuracy based on one repetition of the m-sequence when excluding the first 150 ms. **(e)** Time needed to train the classifiers on all data of each subject. **(f)** Virtual ITR achieved when factoring in 0.5 seconds for gaze shifting.



150 ms. Additionally, the negative correlation between increase in accuracy and number of selected channels can be explained by the fact that the dimensions of the spatiotemporal beamformer increase linearly with the number of channels, thereby requiring more data to accurately estimate the covariance matrix<sup>56,57</sup> (cf., the curse of dimensionality).

## Conclusion

In this study, we have shown that a classifier based on spatiotemporal beamforming is able to accurately discriminate targets encoded by an m-sequence, and could be employed in the context of cVEP BCI. We compared the traditional 60 Hz and the faster 120 Hz stimulus presentation rates, and found that the latter yields more accurate results for equal stimulation lengths, as the encoding sequence can be presented twice as many times as with the 60 Hz case. The maximal median ITR for both stimulus presentation rates and for two iterations of the m-sequence was 100.46 bits/min for the 60 Hz (stimulation time = 2.1 seconds) and 172.87 bits/min for the 120 Hz case rate (stimulation time = 1.05 seconds). We additionally described a transition effect following the onset of the stimulation, similar to SSVEP, and showed that removing the initial 150 ms of the epochs significantly improves classification accuracy when relying on only one repetition of the encoding sequence.

## References

- Nicolas-Alonso, L. F. & Gomez-Gil, J. Brain computer interfaces, a review. *Sensors* **12**, 1211–1279 (2012).
- Bin, G., Gao, X., Wang, Y., Hong, B. & Gao, S. Vep-based brain-computer interfaces: time, frequency, and code modulations [research frontier. *IEEE Computational Intelligence Magazine* **4**, 22–26, <https://doi.org/10.1109/mci.2009.934562> (2009).
- Sutter, E. E. The brain response interface: communication through visually-induced electrical brain responses. *Journal of Microcomputer Applications* **15**, 31–45, [https://doi.org/10.1016/0745-7138\(92\)90045-7](https://doi.org/10.1016/0745-7138(92)90045-7) (1992).
- Hanagata, J. & Momose, K. A method for detecting gazed target using visual evoked potentials elicited by pseudorandom stimuli. In *Proc. 5th Asia Pacific Conf. Medical and Biological Engineering and 11th Int. Conf. Biomedical Engineering (ICBME)* (2002).
- Zierler, N. Linear recurring sequences. *Journal of the Society for Industrial and Applied Mathematics* **7**, 31–48 (1959).
- Golomb, S. W. *et al. Shift register sequences* (Aegean Park Press, 1982).
- Buračas, G. T. & Boynton, G. M. Efficient design of event-related fmri experiments using m-sequences. *NeuroImage* **16**, 801–813, <https://doi.org/10.1006/nimg.2002.1116> (2002).
- Sachs, J., Herrmann, R., Kmec, M., Helbig, M. & Schilling, K. Recent advances and applications of m-sequence based ultra-wideband sensors. 2007 IEEE International Conference on Ultra-Wideband, <https://doi.org/10.1109/icuwb.2007.4380914> (2007).
- Nakanishi, M. & Mitsukura, Y. Periodicity detection for bci based on periodic code modulation visual evoked potentials. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, 665–668 (IEEE, 2012).
- Aminaka, D., Makino, S. & Rutkowski, T. M. Classification accuracy improvement of chromatic and high-frequency code-modulated visual evoked potential-based bci. In *International Conference on Brain Informatics and Health*, 232–241 (Springer, 2015).
- Aminaka, D., Makino, S. & Rutkowski, T. M. Svm classification study of code-modulated visual evoked potentials. In *Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2015 Asia-Pacific*, 1065–1070 (IEEE, 2015).
- Aminaka, D., Makino, S. & Rutkowski, T. M. Eeg filtering optimization for code-modulated chromatic visual evoked potential-based brain-computer interface. In *Symbiotic Interaction*, 1–6 (Springer, 2015).
- Aminaka, D., Makino, S. & Rutkowski, T. M. Chromatic and high-frequency cvep-based bci paradigm. In *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, 1906–1909 (IEEE, 2015).
- Gao, S., Wang, Y., Gao, X. & Hong, B. Visual and auditory brain computer interfaces. *IEEE Transactions on Biomedical Engineering* **61**, 1436–1447, <https://doi.org/10.1109/tbme.2014.2300164> (2014).
- Bin, G. *et al.* A high-speed bci based on code modulation vep. *Journal of Neural Engineering* **8**, 025015 (2011).
- Wei, Q., Feng, S. & Lu, Z. Stimulus specificity of brain-computer interfaces based on code modulation visual evoked potentials. *PLoS one* **11**, e0156416 (2016).
- Spüler, M., Rosenstiel, W. & Bogdan, M. Online adaptation of a c-vep brain-computer interface(bci) based on error-related potentials and unsupervised learning. *PLoS ONE* **7**, e51077, <https://doi.org/10.1371/journal.pone.0051077> (2012).
- Kapeller, C. *et al.* A bci using vep for continuous control of a mobile robot. *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, <https://doi.org/10.1109/embc.2013.6610734> (2013).
- Riechmann, H., Finke, A. & Ritter, H. Using a cvep-based brain-computer interface to control a virtual agent. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **24**, 692–699, <https://doi.org/10.1109/tnsre.2015.2490621> (2016).
- Kapeller, C. *et al.* An electrocorticographic bci using code-based vep for control in video applications: a single-subject study. *Frontiers in Systems Neuroscience* **8**, <https://doi.org/10.3389/fnsys.2014.00139> (2014).
- van Vliet, M. *et al.* Single-trial erp component analysis using a spatiotemporal lcmv beamformer. *IEEE Transactions on Biomedical Engineering* **63**, 55–66, <https://doi.org/10.1109/tbme.2015.2468588> (2016).
- Wittevrongel, B. & Van Hulle, M. M. Faster p300 classifier training using spatiotemporal beamforming. *International Journal of Neural Systems* **26**, 1650014, <https://doi.org/10.1142/s0129065716500143> (2016).
- Luo, A. & Sullivan, T. J. A user-friendly ssvep-based brain-computer interface using a time-domain classifier. *Journal of neural engineering* **7**, 026010 (2010).
- Manyakov, N. V., Chumerin, N., Combaz, A., Robben, A. & Van Hulle, M. M. Decoding ssvep responses using time domain classification. In *IJCCI (ICFC-ICNC)*, 376–380 (2010).
- Wittevrongel, B. & Van Hulle, M. M. Frequency- and phase encoded ssvep using spatiotemporal beamforming. *PLOS ONE* **11**, e0159988, <https://doi.org/10.1371/journal.pone.0159988> (2016).
- Wittevrongel, B. & Van Hulle, M. M. Hierarchical online ssvep spelling achieved with spatiotemporal beamforming. *2016 IEEE Statistical Signal Processing Workshop (SSP)* <https://doi.org/10.1109/ssp.2016.7551800> (2016).
- Brainard, D. H. & Vision, S. The psychophysics toolbox. *Spatial vision* **10**, 433–436 (1997).
- Pelli, D. G. The videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision* **10**, 437–442 (1997).
- Kleiner, M. *et al.* What's new in psychtoolbox-3. *Perception* **36**, 1 (2007).
- Souden, M., Benesty, J. & Affes, S. A study of the lcmv and mvdr noise reduction filters. *IEEE Transactions on Signal Processing* **58**, 4925–4935 (2010).
- Mu, P., Li, D., Yin, Q. & Guo, W. Robust mvdr beamforming based on covariance matrix reconstruction. *Science China Information Sciences* 1–12 (2013).
- Treder, M. S., Porbadnigk, A. K., Avarvand, F. S., Müller, K.-R. & Blankertz, B. The lda beamformer: Optimal estimation of erp source time series using linear discriminant analysis. *NeuroImage* **129**, 279–291 (2016).
- Van Hoey, G. *et al.* Beamforming techniques applied in eeg source analysis. *Proc. ProRISC99* **10**, 545–549 (1999).

34. Belardinelli, P., Ortiz, E. & Braun, C. Source activity correlation effects on lcmv beamformers in a realistic measurement environment. *Computational and mathematical methods in medicine* **2012** (2012).
35. Hong, J. H., Ahn, M., Kim, K. & Jun, S. C. Localization of coherent sources by simultaneous meg and eeg beamformer. *Medical & biological engineering & computing* **51**, 1121–1135 (2013).
36. Van Veen, B. D., Van Drongelen, W., Yuchtman, M. & Suzuki, A. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Transactions on biomedical engineering* **44**, 867–880 (1997).
37. Bewick, V., Cheek, L. & Ball, J. Statistics review 13: receiver operating characteristic curves. *Critical care* **8**, 508 (2004).
38. Lasko, T. A., Bhagwat, J. G., Zou, K. H. & Ohno-Machado, L. The use of receiver operating characteristic curves in biomedical informatics. *Journal of biomedical informatics* **38**, 404–415 (2005).
39. Vapnik, V. N. & Vapnik, V. *Statistical learning theory*, vol. 1 (Wiley New York, 1998).
40. Hsu, C.-W., Chang, C.-C., Lin, C.-J. *et al.* A practical guide to support vector classification (2003).
41. Keerthi, S. S. & DeCoste, D. A modified finite newton method for fast solution of large scale linear svms. *In Journal of Machine Learning Research* **6**, 341–361 (2005).
42. Combaz, A. *et al.* A comparison of two spelling brain-computer interfaces based on visual p3 and ssvep in locked-in syndrome. *PloS one* **8**, e73691 (2013).
43. Combaz, A. *et al.* Towards the detection of error-related potentials and its integration in the context of a p300 speller brain-computer interface. *Neurocomputing* **80**, 73–82 (2012).
44. Di Russo, F. & Spinelli, D. Electrophysiological evidence for an early attentional mechanism in visual processing in humans. *Vision research* **39**, 2975–2985 (1999).
45. Di Russo, F., Teder-Sälejärvi, W. A. & Hillyard, S. A. Steady-state vep and attentional visual processing. The cognitive electrophysiology of mind and brain (Zani A, Proverbio AM, eds) 259–274 (2002).
46. Nakanishi, M., Wang, Y., Wang, Y.-T., Mitsukura, Y. & Jung, T.-P. A high-speed brain speller using steady-state visual evoked potentials. *International journal of neural systems* **24**, 1450019 (2014).
47. Wolpaw, J. R., Ramoser, H., McFarland, D. J. & Pfurtscheller, G. Eeg-based communication: improved accuracy by response verification. *IEEE transactions on Rehabilitation Engineering* **6**, 326–333 (1998).
48. Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G. & Vaughan, T. M. Brain-computer interfaces for communication and control. *Clinical neurophysiology* **113**, 767–791 (2002).
49. Volosyak, I., Valbuena, D., Luth, T. & Gräser, A. Towards an ssvep based bci with high itr. *IEEE Trans. Biomed. Eng.* (2010).
50. Chen, X., Chen, Z., Gao, S. & Gao, X. A high-itr ssvep-based bci speller. *Brain-Computer Interfaces* **1**, 181–191 (2014).
51. Lin, K., Chen, X., Huang, X., Ding, Q. & Gao, X. A hybrid bci speller based on the combination of emg envelopes and ssvep. *In Applied informatics*, vol. 2, 1 (Springer Berlin Heidelberg, 2015).
52. Chen, X. *et al.* High-speed spelling with a noninvasive brain-computer interface. *Proceedings of the national academy of sciences* **112**, E6058–E6067 (2015).
53. Kübler, A., Neumann, N., Wilhelm, B., Hinterberger, T. & Birbaumer, N. Predictability of brain-computer communication. *Journal of Psychophysiology* **18**, 121–129 (2004).
54. Kübler, A. & Birbaumer, N. Brain-computer interfaces and communication in paralysis: Extinction of goal directed thinking in completely paralysed patients? *Clinical neurophysiology* **119**, 2658–2666 (2008).
55. Brunner, C., Allison, B., Altstätter, C. & Neuper, C. A comparison of three brain-computer interfaces based on event-related desynchronization, steady state visual evoked potentials, or a hybrid approach using both signals. *Journal of neural engineering* **8**, 025010 (2011).
56. Pruzek, R. M. High dimensional covariance estimation: Avoiding the ‘curse of dimensionality’. *In Proceedings of the First US/Japan Conference on the Frontiers of Statistical Modeling: An Informational Approach*, 233–253 (Springer, 1994).
57. Schoukens, J. & Pintelon, R. *Identification of linear systems: a practical guideline to accurate modeling* (Elsevier, 2014).

## Acknowledgements

B.W. is supported by a Strategic Basic Research (SBO) grant, funded by VLAIO (Flemish Agency for Innovation and Entrepreneurship). M.M.V.H. is supported by research grants received from the Financing program (PFV/10/008), an interdisciplinary research project (IDO/12/007), and an industrial research fund project (IOF/HB/12/021) of the KU Leuven, the Belgian Fund for Scientific Research – Flanders (G088314N, G0A0914N), the Inter-university Attraction Poles Programme – Belgian Science Policy (IUAP P7/11), the Flemish Regional Ministry of Education (Belgium) (GOA 10/019), and the Hercules Foundation (AKUL 043).

## Author Contributions

B.W. conceived the experiment, B.W. conducted the experiments, B.W. and E.V. analysed the results. All authors wrote and reviewed the manuscript.

## Additional Information

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017