An Event-Driven AR-Process Model for EEG-Based BCIs With Rapid Trial Sequences

Paula Gonzalez-Navarro, Yeganeh M. Marghi, Bahar Azari, Murat Akçakaya, and Deniz Erdoğmuş

***Abstract*— Electroencephalography (EEG) is an effective non-invasive measurement method to infer user intent in brain-computer interface (BCI) systems for control and communication, however, these systems often lack sufficient accuracy and speed due to low separability of class-conditional EEG feature distributions. Many factors impact system performance, including inadequate training datasets and models' ignorance of the temporal dependency of brain responses to serial stimuli. Here, we propose a signal model for event-related responses in the EEG evoked with a rapid sequence of stimuli in BCI applications. The model describes the EEG as a superposition of impulse responses time-locked to stimuli corrupted with an autoregressive noise process. The performance of the signal model is assessed in the context of RSVP keyboard, a language-model-assisted EEG-based BCI for typing. EEG data obtained for model calibration from 10 healthy participants are used to fit and compare two models: the proposed sequence-based EEG model and the trial-based feature-class-conditional distribution model that ignores temporal dependencies, which has been used in the previous work. The simulation studies indicate that the earlier model that ignores temporal dependencies may be causing drastic reductions in achievable information transfer rate (ITR). Furthermore, the proposed model, with better regularization, may achieve improved accuracy with fewer calibration data samples, potentially helping to reduce calibration time. Specifically, results show an average 8.6% increase in (cross-validated) calibration AUC for a single channel of EEG, and 54% increase in the ITR in a typing task.**

***Index Terms*— EEG, BCI, signal model, event-related potential, RSVP keyboard.**

I. Introduction

ELECTROENCEPHALOGRAM (EEG)-based brain-computer interfaces (BCIs) offer alternative communication channels for human computer interaction. This technology can be especially important for individuals with severe speech or motor disorders who cannot use communication pathways designed for healthy individuals or other assistive technology input modalities. By detecting specific EEG patterns, brain activities can be translated into actions to accomplish a task.

ERPs are commonly employed in EEG-based BCIs to determine user intent in various BCI paradigms, such as motor imagery, speller, etc [1]–[7]. To increase accuracy, repeated stimuli are often used at the cost of increasing time-to-decision. To tackle the problem of low accuracy while maintaining the appropriate time-to-decision, the error-related potentials (ErrP) has been recently proposed in various BCI setting. For instance, in BCI settings with auditory stimuli where accuracy is relatively low incorporation of ErrPs improves the overall performance of the system [8]. In addition, the presence of ErrP in cognitive task-based BCI [9] and speller BCI [10] has shown significant improvement. In the context of speller BCIs, these were previously studied as feedback-related potentials (FRPs) [11]–[14]. Fusion of EEG evidence from ERP and ErrP/FRP trials along with context priors from a language model, coupled with active learning techniques, can significantly improve the speed and accuracy of BCIs that rely on ERPs [15], [16].

In ERP-driven BCI designs, EEG feature models typically rely on feature extraction and modeling of feature distributions conditioned on intent. These models typically assume that features for individual trials are statistically independent conditioned on the intent; they also typically assume Gaussianity for EEG features either explicitly or implicitly, with covariances that are approximately equal (which amounts to assuming a wide sense stationary background noise) [17], [18]. In particular, the assumption of EEG features being conditionally independent given intent is usually incorrect due to the rapid succession of stimuli and overlapping windows in time over which these features are extracted.

This study proposes an EEG signal model that represents EEG signal in each channel as a linear combination of evoked brain activities that are driven by all stimuli and other events in the paradigm corrupted by colored background noise that is characterized by an autoregressive (AR) Gaussian process model. This model explicitly attempts to capture the underlying statistical behavior of EEG in response to sequences of stimuli (events) in EEG-based BCIs, particularly for rapid stimulus sequences such as typing-based or auditory BCIs. The proposed signal model can enhance the BCI system performance, i.e. classification accuracy and information transfer rate by improving assessment of temporal dependencies across samples.

Manuscript received July 2, 2018; revised December 20, 2018 and February 17, 2019; accepted March 3, 2019. Date of publication March 8, 2019; date of current version May 9, 2019. This work was supported in part by NSF under Grant IIS-1149570, Grant CNS-1544895, Grant IIS-1715858, Grant IIS-1717654, and Grant IIS-1844885, in part by DHHS under Grant 90RE5017-02-01, and in part by NIH under Grant R01DC009834. *(Paula Gonzalez-Navarro and Yeganeh M. Marghi are co-first authors.) (Corresponding author: Paula Gonzalez-Navarro.)*

P. Gonzalez-Navarro, Y. M. Marghi, B. Azari, and D. Erdoğmuş are with the Cognitive Systems Laboratory, Electrical and Computer Engineering Department, Northeastern University, Boston, MA 02115 USA (e-mail: gonzaleznavarro@ece.neu.edu).

M. Akçakaya is with the Department of Electrical and Computer Engineering, University of Pittsburgh, PA 15261 USA.

Digital Object Identifier 10.1109/TNSRE.2019.2903840

II. Signal Model for EEG-Based BCIs

In BCIs that rely on ERPs there are typically multiple types of events that may evoke a brain response and these events may happen in rapid succession. In this paper, for illustration purposes, we will focus on designs that employ a combination of ERPs in response to rapidly presented stimuli as well as FRPs in response to isolated stimuli. The ERP-inducing sequences will consist primarily of negative intent trials, while rare trials that evoke positive intent will occur intermittently. The following isolated FRP trials will tend to evoke associated positive or negative intent responses in a more balanced fashion across these two class labels. The proposed signal model that considers the entire sequence of stimuli will be referred to as a sequence-based (SB) model and is expected to accurately model EEG statistics with only a few parameters.

A. Sequence-Based EEG Model

To enhance the probabilistic representation of EEG measurements, the temporal dependency between EEG trials are considered as a parametric temporal structure over the EEG sequences. An EEG channel signal is modeled as a linear combination of three time series components: (1) task-dependent brain activities, e.g., visual evoked potential (VEP), ERP, or FRP, (2) task-independent brain activities, i.e., underlying cortical activities, and (3) an additive measurement noise [19]. The task-dependent brain activities in response to a single or series of stimuli can be described as a linear combination of different temporal activities, such as initial sensory responses to a stimulus, task-dependent perceptual processes, decision making, and post-response evaluation [20]. By assuming the EEG sequence as a stationary Gaussian process, it can be modeled as a time series that consists of five components: three temporal filters corresponding to the task-dependent brain activities for ERP and FRP sequences, an AR process representing task-independent brain activities, and the measurement noise. Accordingly, a single-channel EEG signal y at the j-th sequence and time sample n can be expressed as:

$$y^j[n] = \sum_{t \in T^j} h[n] * \delta[n-t] + s[n] + \epsilon[n], \tag{1}$$

where $h[n]$ is the impulse response function of a temporal filter that describes the dynamic of the stimulus-dependent brain activities, t is the onset of each stimulus and T^j is a set including the onset of all stimuli at the j-th sequence. The variable $\epsilon[n]$ is the additive measurement noise with $\mathcal{N}(0, \sigma_\epsilon^2)$ and $s[n]$ is the AR process exhibiting temporal correlation expressed as:

$$s[n] = \sum_{k=1}^{K} a_k s[n-k] + \omega[n], \tag{2}$$

where $\omega[n]$ is a Gaussian process with distribution $\mathcal{N}(0, \sigma_\omega^2)$. The stimulus-dependent brain activities is the direct response to a particular sensory (visual, auditory) stimuli, cognitive (memory, inhibition) process, or motor command [20]. Figure 1 shows the well-studied brain waveforms including

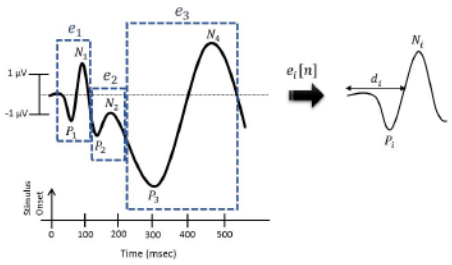

Fig. 1. Time-locked electrical activities of the human brain in response to a particular stimulus. The entire brain activities can be described as a linear combination of delayed negative and positive neural deflections shown in each box, i.e., e_1, e_2, and e_3.

$N_{1,2,4}$ and $P_{1,2,3}$ in response to various types of stimuli that are split into e_i waves. Therefore, temporal dynamics of the impulse response, $h[n]$, can be described as a linear combination of negative and positive neural deflections after a stimulus. According to the latency and amplitude of each pair of positive and negative peaks, we decompose $h[n]$ into three waves as follows:

$$h[n] = \sum_{i=1}^{3} e_i[n - d_i], \tag{3}$$

where d_i is the delay of each wave. The e_1 wave that includes N_1 and P_1 is associated with the initial sensory response, e_2 includes N_2 and P_2 and is more related to the decision making and categorization process [20], and lastly, e_3 consists of P_3 and N_4 and corresponds to the infrequent target stimulus [21] and post-response evaluation [20]. Based on the temporal structure of each e_i component, this wave can be modeled by an exponential polynomial function:

$$e_i[n] = f(n, \tau_i, L_i) \tag{4}$$

$$f(n, \tau_i, L_i) = \sum_{l=1}^{L_i} b_{il} \left(\frac{n}{l}\right)^l \exp(-n/\tau_i) u[n]$$

$$= \mathbf{g}_{\tau_i, L_i}^T[n] \mathbf{b}_i \tag{5}$$

Here, L_i represents the order of polynomial terms and b_i expresses the amplitude of each exponential term. In the simplified vector form, $\mathbf{b}_i = \left[b_{i,1}\ b_{i,2}\ \ldots\ b_{i,L_i}\right]^T$ and $\mathbf{g}_{\tau_i, L_i}[n] = \left[n\ (n/2)^2\ \ldots\ (n/L_i)^{L_i}\right]^T \exp(-n/\tau_i) u[n]$.

Accordingly, (1) can be rewritten as:

$$y^j[n] = \sum_{t \in T^j} e_1[n] * \delta[n - t - d_1] + e_2[n] * \delta[n - t_*^j - d_2]$$
$$e_3[n] * \delta[n - t_*^j - d_3] + s[n] + \epsilon[n], \tag{6}$$

where t_*^j is the onset corresponding to the target stimulus. Lastly, (6) can be expressed in vector form for all samples in sequence j as:

$$\mathbf{y}^j = \mathbf{X}^j \boldsymbol{\beta} + \mathbf{z}, \tag{7}$$

where $\mathbf{y}^j \in \mathbb{R}^{(N-K)\times 1}$, K is the order of the AR process and N is the number of samples in a sequence. Accordingly, the SB model in (7), includes brain responses to the entire

trials in a sequence with the overlap between target, non-target, and independent task activities. In fact, this model enables us to properly decouple different brain activities as parametric components. $\mathbf{X}^j \in \mathbb{R}^{(N-K)\times(L_1+L_2+L_3)}$ is called the design matrix, which includes the combination of all exponential polynomial waves according to the onset of the stimuli and the delays.

$$\mathbf{X}^j = \left[\mathcal{X}_1^j \ \mathcal{X}_2^j \ \mathcal{X}_3^j\right], \tag{8}$$

where each $\mathcal{X}_i^j$ matrix is defined as:

$$\mathcal{X}_1^j[n] = \sum_{t\in T^j} \mathbf{g}_{\tau_1,L_1}^T[n-t-d_1]$$

$$\mathcal{X}_2^j[n] = \mathbf{g}_{\tau_2,L_2}^T[n-t_*^j-d_2], \quad \mathcal{X}_3^j[n] = \mathbf{g}_{\tau_3,L_3}^T[n-t_*^j-d_3] \tag{9}$$

The amplitude parameter $\boldsymbol{\beta} \in \mathbb{R}^{(L_1+L_2+L_3)\times 1}$ is also defined as:

$$\boldsymbol{\beta} = \begin{bmatrix}\mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_3\end{bmatrix}^T. \tag{10}$$

In (7), $\mathbf{z}$ includes the AR(K) process plus measurement noise, as:

$$\begin{aligned}\mathbf{z} &= \mathbf{s} + \boldsymbol{\epsilon}\\ \mathbf{s} &= \tilde{\mathbf{S}}\mathbf{a}_K + \boldsymbol{\omega}\end{aligned} \tag{11}$$

where $\mathbf{a}_K = \begin{bmatrix}a_1 \ a_2 \ \dots \ a_K\end{bmatrix}^T$, $\mathbf{s} \in \mathbb{R}^{(N-K)\times 1}$, and $\tilde{\mathbf{S}} \in \mathbb{R}^{(N-K)\times K}$. Each row of matrix $\tilde{\mathbf{S}}$ is defined as $\tilde{\mathbf{s}}[n] = \begin{bmatrix}s[n] \ s[n-1] \ \dots \ s[n-K+1]\end{bmatrix}$.

The AR(K) process in (7) can be simplified as:

$$\mathbf{z} = \tilde{\mathbf{S}}\mathbf{a}_K + \boldsymbol{\xi}, \tag{12}$$

where $\boldsymbol{\xi}$ is a white Gaussian process with $\mathcal{N}(\mathbf{0}, \sigma_\xi^2\mathbf{I})$ and $\sigma_\xi^2 = \sigma_\epsilon^2 + \sigma_\omega^2$. Accordingly, the probability distribution of an EEG sequence (ERP or FRP) can be expressed as a Gaussian process $\mathbf{y} \sim \mathcal{N}(\mathbf{X}\boldsymbol{\beta}, \boldsymbol{\Sigma})$, in which the covariance matrix $\boldsymbol{\Sigma}$ has a Toeplitz structure as a function of parameters $\mathbf{a}_K$ and σ_ξ^2.

Since our goal is to model two sets of EEG sequences, i.e., $\mathbf{y}_{erp}$ and $\mathbf{y}_{frp}$, the model is required to learn two sets of parameters for ERP and FRP sequences. Moreover, due to the fact that the FRP sequences only include a single stimulus that generates either a positive or a negative response, $\boldsymbol{\beta}$ will be different for each response type. Therefore, the model parameters for ERP sequences can be defined as $\boldsymbol{\theta}_{erp} = [\mathbf{a}_K, \boldsymbol{\beta}_{erp}, \sigma_\xi^2]$ and for FRP sequences as $\boldsymbol{\theta}_{frp} = [\mathbf{a}_K, \boldsymbol{\beta}_{+frp}, \boldsymbol{\beta}_{-frp}, \sigma_\xi^2]$. The maximum likelihood estimation (MLE) method has been used for $\boldsymbol{\theta}_{erp}$ and $\boldsymbol{\theta}_{frp}$ estimation by solving the following optimization problem:

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \ p(\mathbb{Y} \mid \mathbb{X}, \boldsymbol{\theta}). \tag{13}$$

Here, $\mathbb{Y} = \{\mathbf{y}^1, \mathbf{y}^2, \dots, \mathbf{y}^{N_s}\}$ and $\mathbb{X} = \{\mathbf{X}^1, \mathbf{X}^2, \dots, \mathbf{X}^{N_s}\}$ are sets of supervised EEG sequences and design matrices for N_s number of sequences. Assuming that EEG sequences are independently distributed, the loglikelihood function can be written as:

$$\begin{aligned}\log p(\mathbb{Y} \mid \boldsymbol{\theta}) = {} & N_s(N-K)\log(2\pi) + N_s \log(\mid \boldsymbol{\Sigma} \mid^{-1})\\ & + \sum_{j=1}^{N_s} \left(\mathbf{y}^j - \mathbf{X}^j\boldsymbol{\beta}\right)^T \boldsymbol{\Sigma}^{-1} \left(\mathbf{y}^j - \mathbf{X}^j\boldsymbol{\beta}\right).\end{aligned} \tag{14}$$

In order to jointly estimate all parameters, the cyclic descent algorithm [19] is used to solve (13). The other parameters of the model $(L_i, \tau_i, d_i), i = \{1, 2, 3\}$ are learned as hyperparameters during cross validation. The AR order K is learned using the Bayesian information criterion (BIC).

B. User Intent Inference

After learning the model parameters, a maximum a-posteriori (MAP) method is utilized to estimate the user intent (target stimulus or correct and incorrect stimulus in a sequence). The posterior probability distribution of a stimulus is updated according to the observed EEG sequences and the context information (prior). The posterior update process following a sequential procedure that updates the posterior probability after collecting an ERP or FRP sequence. This sequential process can be performed using either a single- or a multi-channel EEG model.

1) Using Single EEG Channel: We estimate the user intent that maximizes the probability distribution over the EEG sequence for a given design matrix that is a function of the onset of stimuli and the model parameters and the non-EEG contextual evidence. Assuming $\psi \in \mathcal{D}$ as all the potential candidates for user intent ψ^*, where $\mathcal{D}$ is the dictionary comprising of all such candidates, the posterior probability can be estimated using MAP as follows:

$$\begin{aligned}\hat{\psi} &= \arg\max_{\psi\in\mathcal{D}} \ p\left(\psi^* = \psi \mid \mathbb{Y}, \boldsymbol{\theta}, C\right)\\ &= \arg\max_{\psi\in\mathcal{D}} \ p\left(\mathbb{Y} \mid \psi, \boldsymbol{\theta}\right) p(\psi^* = \psi \mid C),\end{aligned} \tag{15}$$

where $\mathbb{Y}$ represents obtained EEG sequences that can include both ERP or combined ERP and FRP, and C represents the non-EEG contextual evidence defined according to the experimental design. Since the design matrix includes the order and timing information of all stimuli within a sequence, we can say that $p\left(\mathbb{Y} \mid \psi, \boldsymbol{\theta}\right) = p\left(\mathbb{Y} \mid \mathbb{X}(\psi), \boldsymbol{\theta}\right)$. Assuming ERP and FRP sequences as independent EEG sequences, the likelihood of the collected combined ERP/FRP sequences can be expressed as:

$$\begin{aligned}p\left(\mathbb{Y} \mid \mathbb{X}(\psi), \boldsymbol{\theta}\right) = {} & p\left(\mathbb{Y}_{erp} \mid \mathbb{X}_{erp}(\psi), \boldsymbol{\theta}_{erp}\right)\\ & p\left(\mathbb{Y}_{frp} \mid \mathbb{X}_{frp}(\psi), \boldsymbol{\theta}_{frp}\right)\end{aligned} \tag{16}$$

According to (17) and (16), the sequence-based (SB) model represents the entire EEG signal obtained in response to the sequence of trials, unlike the traditional trial-based (TB) method that assumes trials within a sequence are independent events.

2) Using Multiple EEG Channels: Assuming EEG channels are independent attributes, weighted naive Bayes (WNB) can be used to update the posterior probability as following:

$$\hat{\psi} = \arg\max_{\psi \in \mathcal{D}} \prod_{c=1}^{N_c} p\left(\mathbb{Y}_c \mid \mathbb{X}(\psi), \theta_c\right)^{w_c} p(\psi^* = \psi \mid C), \quad (17)$$

where c is the channel index, N_c is the total number of channels, and w_c denotes the weight of the likelihood score of each channel. Here, we used a regularized least squares (RLS) method to estimate channel weights in order to improve the inference performance.

Applying the proposed temporal structure and using sequence-based inference for both FRP and ERP sequences can result in more precise user intent estimation in comparison with the trial-based approach, which will be demonstrated in the following sections.

III. Experiment: An EEG-Based BCI Typing System

To assess the SB EEG model, we used a language-model-assisted EEG-based BCI typing system called RSVP Keyboard [3]. This system utilizes RSVP paradigm which stands for Rapid Serial Visual Presentation [2]. In RSVP Keyboard a set of pseudo-randomly ordered symbols are rapidly presented as a time series on a prefixed location on the screen in a pseudo-random order to evoke visual potentials [4], [5]. The system uses a finite set of 28 symbols from the English alphabet, in addition to symbols for space and backspace; that is $\mathcal{D} = \{A, B, C, \ldots, Z, \_, <\}$. Each flashing symbol is defined as an EEG *trial*, and a set of trials that have been presented with a small time gap in between (e.g. 200 ms) is called an EEG *sequence*. Each sequence is appended with the top candidate in the alphabet *prospect symbol* as feedback according to the posterior probability after ERP/FRP sequence fusion. This prospect symbol is presented as a single flash as an RSVP trial in a different color than the regular trials to induce positive or negative FRP in the EEG response depending on the prospect's correctness. Due to low signal to noise ratio (SNR) of EEG, the system is usually required to query the user with more than one sequence and prospect symbols to achieve a desired confidence level. The set of sequences and prospect symbols which lead to a decision is called an *epoch*. It is assumed that the target symbol remains unchanged at each epoch. Figure 2 represents a schematic of EEG sequences in the RSVP Keyboard including a series of of visual stimuli (alphabet) in an ERP sequence and a feedback stimulus as an FRP sequence. In this example FRP sequence corresponds to a positive response, because it matches with the target symbol at trial 6; the rest of the trials correspond to non-target symbols.

The RSVP Keyboard has three important system operation modes that are utilized in this study:
(1) *Calibration*: The calibration is used to learn the class-conditional distribution of the ERP (target vs. non-target symbols) and FRP sequences (correct vs. incorrect symbols). A user is asked to focus on a set of predefined target symbols shown on the screen, to record labeled EEG data.

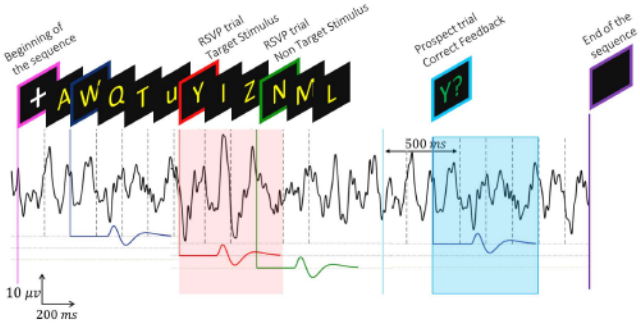

Fig. 2. Schematic of a sequence for the RSVP Keyboard framework. A series of symbols including non-target and target symbols is shown at a prefixed position on the screen consecutively over time in rapid trial fashion. The sequence starts with a + symbol, which is indicated by the magenta line, the ending of the RSVP sequence, a blank screen, is shown with a blue line, and the ending of the sequence, a blank screen, is shown with a purple line. At the end of the sequence, a prospect symbol is appended.

(2) *Copy Phrase*: This mode is designed to assess the system and the user performance in terms of speed and accuracy in the presence of a language model. The user is given a set of predefined phrases. Each phrase includes a missing word and the user is asked to complete the word.

(3) *Simulation*: This mode is base on a proposed probabilistic simulation framework in [22]. A Copy Phrase task is simulated without user intervention. We use Monte-Carlo sampling to draw samples from class distributions using the class conditional distributions learned from Calibration mode. This mode reports the system performance in terms of speed and accuracy.

Detailed information about RSVP Keyboard can be found in [3], [5].

IV. Experimental Results and Discussions

EEG data was collected from 10 healthy participants while they performed the RSVP Keyboard task. All participants had consented to participate according to the protocol (IRB-130107) approved by the IRB of Northeastern University. In our typing experiment, each sequence included 14 trials and single-trial feedback. The EEG signals were acquired using a g.USBamp biosignal amplifier with active g.Butterfly electrodes at a sampling rate of 256 Hz, from 16 EEG sites: Fp1, Fp2, F3, F4, Fz, Fc1, Fc2, Cz, P1, P2, C1, C2, Cp3, Cp4, P5 and P6. To improve the SNR, and to eliminate drifts, signals were filtered by an FIR linear-phase bandpass ([1.5, 42] Hz) filter. In particular, in these experiments we used the calibration mode of the system explained in section III. The following subsections describe several studies to analyze and assess the performance of the proposed SB signal model, compared to a trial-based model.

In trial-based models, it is typically assumed that EEG trials are independent observations coming from a Gaussian process. For these cases, regularization [3] or spatio-temporal structuration [23] are known techniques to estimate the class-dependent mean and covariance. Other than ignoring the dependency among trials, these approaches suffer from large number of parameters required in estimating the mean and

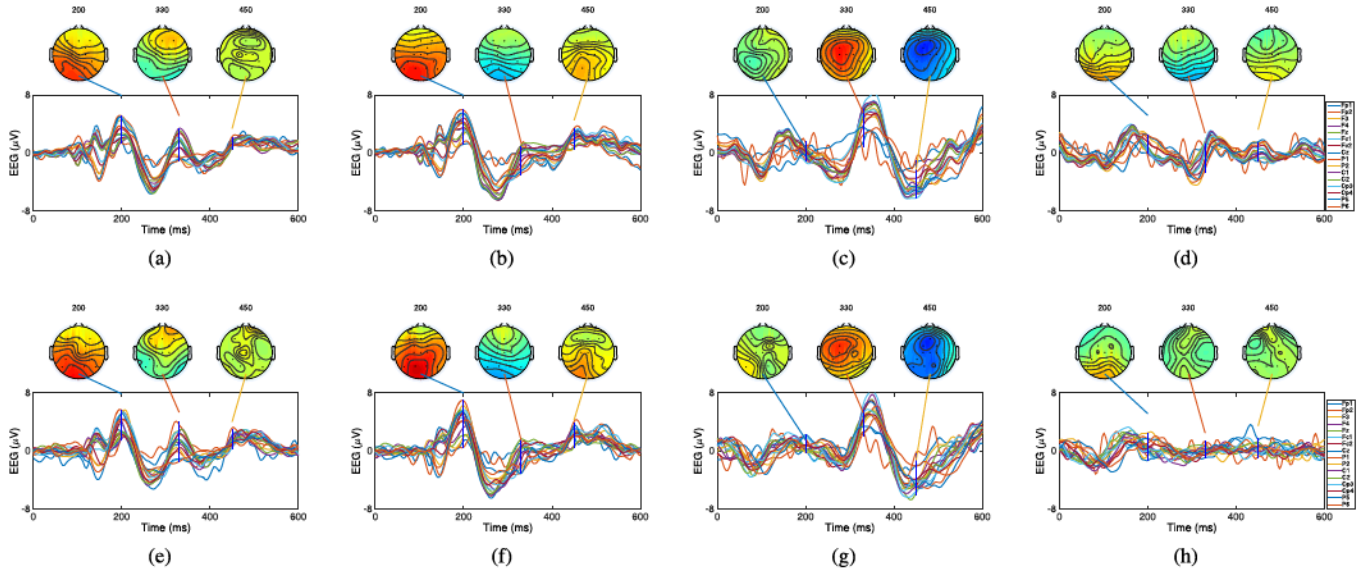

Fig. 3. Performance of the proposed signal model in reflecting the time-series characteristics of the EEG signal. First row average of 16 original time-locked EEG signal for the (a) incorrect feedback, (b) correct feedback, (c) target events (d) target events. Second row average of 16 estimated time-locked EEG signal for the (e) incorrect feedback, (f) correct feedback (g) target events, and (h) target events. Topo-plots above each panel show the average topography of the entire channels at three different time points: 200, 330, and 450 ms.

covariance. In the TB model used for this section, each EEG trial was defined as a time window of $[0, 500)$ ms of signal after each flash onset. After each trial, the model evaluates acquired ERPs and FRPs in the EEG and assesses their likelihood of being in response to a target letter and their likelihood of being in response to a correct or incorrect letter, respectively. However, the proposed SB model represents the entire EEG signal obtained in response to a sequence of trials, as opposed to modeling the EEG feature vector for each trial assuming independence. It evaluates acquired ERPs and FRPs in the EEG for the entire sequence and assesses their likelihood of being the target letter in different locations of the sequence, corresponding to each symbol presented in a sequence. Evidence from EEG and context information are fused using a naive Bayes assumption to make a joint decision using MAP inference. Here, the non-EEG evidence is provided by an n-gram language model, which estimates the conditional probability of every letter in the alphabet based on $n-1$ previously typed letters in a Markov model framework. The RSVP Keyboard inference model is explained in detail in [3].

A. Goodness of Fit

Qualitative and quantitative measures of goodness of fit will be presented. We conducted a Chi-squared goodness-of-fit test to evaluate the performance of the proposed signal model in capturing the probability distribution of the EEG signal. This performance evaluation can also be performed visually by comparing topographic scalp maps from actual EEG signals with those generated by the SB model.

1) Chi-Squared Goodness-of-Fit Test: A Chi-squared goodness-of-fit test was performed to evaluate the performance of the developed SB model in capturing the underlying distribution of the EEG sequences. This test determines whether independent samples are consistent with a Gaussian distribution. In our case, the sample data comes from a colored Gaussian process, so we used whitened data using the captured distribution from the model. The significance level indicates the estimated probability with which the null hypothesis is rejected. For the ERP and FRP models, the significance levels averaged over 10 participants are 12% and 38%, respectively.

2) Topographic Scalp Maps: Figure 3 illustrates the ability of the signal model to estimate the EEG measurement. The average EEG signal across trials for the target and non-target stimuli and for the correct and incorrect feedback of the signal model can be compared against the actual EEG signal. It can be observed that the signal model powerfully captures the important temporal features of EEG time series.

B. Classification Performance Assessment

As a classification assessment indicator for each model, the area under the receiver operating characteristics (ROC) curve (AUC) are reported in Table I for the both ERP and FRP sequences using TB and SB models. Table I presents the AUCs for all 10 participants in the BCI typing task. AUC values are calculated based on the 10 fold cross validation on the calibration data sets. Same calibration data has been used for both TB and SB models. In the case of TB, the distributions were captured over trials without imposing any temporal structural over the signal [3]. However, for SB, the probability distributions were learned through signal model in (1). It can be seen that for a single channel (Fz), applying the temporal structure can improve the classification performance and the AUC values obtained from the SB model is higher than the TB model. In particular, we observed more improvement for the ERP sequences using the SB model, since it includes the temporal dependency between trials. This improvement is not observed for the FRP sequences. These results suggest that a sequence based signal model can be highly beneficial for capturing the temporal dependency across trials.

TABLE I
CLASSIFICATION RESULTS FOR TB AND SB EEG MODEL, FOR SINGLE AND MULTI EEG CHANNEL. THE REPORTED VALUES ARE THE MEAN AUC ACROSS 10 FOLD CROSS VALIDATION. ALL OF THE SINGLE-CHANNEL RESULTS BELONG TO THE EEG CLASSIFICATION USING THE FZ CHANNEL

	ERP				FRP			
User	TB(Fz)	SB(Fz)	TB(16)	SB(16)	TB(Fz)	SB(Fz)	TB(16)	SB(16)
1	61.8 ± 07.2	71.1 ± 09.5	82.4 ± 05.4	88.9 ± 04.9	77.3 ± 12.4	79.3 ± 08.7	79.5 ± 12.6	81.2 ± 06.8
2	59.9 ± 04.2	69.3 ± 07.2	84.5 ± 05.6	88.3 ± 03.6	58.2 ± 07.4	63.0 ± 10.3	75.4 ± 10.6	79.0 ± 16.2
3	68.6 ± 04.9	82.5 ± 07.2	92.2 ± 04.0	95.7 ± 02.2	76.9 ± 12.9	74.1 ± 09.7	82.8 ± 06.6	82.0 ± 11.6
4	68.1 ± 09.5	79.8 ± 11.0	85.3 ± 05.2	92.0 ± 04.5	83.2 ± 09.8	71.7 ± 12.4	83.9 ± 07.1	77.1 ± 9.7
5	63.0 ± 04.6	69.5 ± 06.1	85.7 ± 02.8	93.0 ± 02.7	81.0 ± 12.6	74.7 ± 06.7	91.5 ± 07.4	87.4 ± 09.4
6	62.1 ± 03.8	64.8 ± 08.6	72.2 ± 03.6	79.5 ± 06.1	62.3 ± 15.8	62.7 ± 12.3	64.9 ± 13.2	69.9 ± 12.6
7	66.1 ± 05.6	73.1 ± 07.3	85.9 ± 03.5	87.5 ± 03.2	76.9 ± 06.9	71.2 ± 05.9	82.5 ± 07.3	77.1 ± 14.0
8	66.6 ± 08.6	76.3 ± 12.9	81.7 ± 10.9	88.2 ± 05.1	53.4 ± 07.9	57.4 ± 09.4	67.1 ± 08.4	66.2 ± 12.8
9	53.1 ± 06.3	51.3 ± 09.5	78.5 ± 04.1	87.5 ± 06.0	62.2 ± 14.0	61.2 ± 14.3	81.9 ± 14.3	70.4 ± 09.4
10	85.7 ± 09.1	74.1 ± 10.7	95.7 ± 06.0	98.1 ± 00.9	77.3 ± 06.3	72.9 ± 07.7	80.6 ± 09.5	82.1 ± 07.8

C. Simulations

In this study, a Copy Phrase task was simulated with the simulation mode of RSVP Keyboard explained in III using EEG data collected during the calibration sessions. For both models, the probability distributions were captured by four EEG signals models: (1) TB ERP, (2) TB ERP-FRP, (3) SB ERP(Fz), (4) SB ERP-FRP. Calibration data consist of real ERP-FRP EEG data from a single-channel (Fz) and multiple channels for 10 healthy participants. Using the probability distributions learned for each model in study IV-B, the probabilistic simulation framework in [22] has been used by taking 100 Monte-Carlo simulations of a Copy Phrase task with 10 predefined sentences. The typing performances for all models are reported by ITR (bit/sequence). This measure summarizes the accuracy and speed into a single metric and is commonly used to measure BCI system performance. Figures 4a and 4b illustrate the average ITR values for single-channel and multiple channel, respectively, for all 10 subjects sorted according to the calibration session AUC values obtained by SB-ERP multiple channel model. The results in Figure 4a show that for single-channel case, the SB model outperforms the TB model for ERP and ERP-FRP sequences. The results presented in Figure 4b show the ITR values for multiple channel for the different signal models. Although the proposed model for multiple-channel does not demonstrate to be as large an improvement as for single-channel, we can see an overall trend of better performance for the SB model compared to the TB model in the accuracy and the speed of the user intent detection.

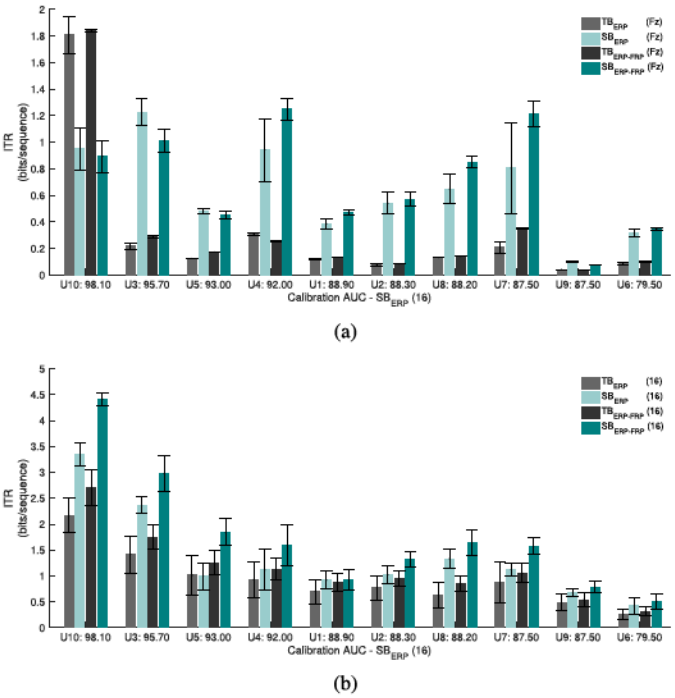

Fig. 4. Average of information transfer rate for 4 signal models with error bars indicating standard deviation: (1) TB model ERP, (2) TB ERP-FRP, (3) SB ERP, and (4) SB ERP-FRP. The results belongs to 100 Monte Carlo simulation of a copy phrase task with 10 predefined sentences, using synthetic EEG features from 4 different models calibrated with real ERP/FRP EEG data (a) form Fz channel (b) from 16 channels locations from 10 healthy participants.

D. Calibration Length Assessment

Calibration length governs the size of the dataset (number of sequences) used in the training process. Figure 5 shows the AUC computed for different length of calibration dataset for training the SB model and the TB model applied to ERP and FRP sequences. The results for single and multiple channel ERP and FRP sequences are illustrated in Figures 5a, 5b, 5c and 5d, respectively. Note that there are 14 trials per ERP sequences for training the model, while there is only one trial per FRP sequences. Since the SB model uses only a few number of parameters, it can learn the signal distribution using less number of samples, while for the TB model the performance worsens quickly when we use a smaller portion of the dataset. Therefore, it can be concluded that the SB model achieves better performance for the FRP sequences with smaller training samples. For the ERP sequences, although the SB model provides higher AUCs, the improvement is not quite significant. The reason for this is that even with a smaller dataset size, we have 14 (trials) times more data in ERP sequences than FRP sequences.

E. Model Complexity

It is worthwhile to note that the proposed signal model has significantly fewer parameters, and thus lower model complexity, than earlier models that consider each trial

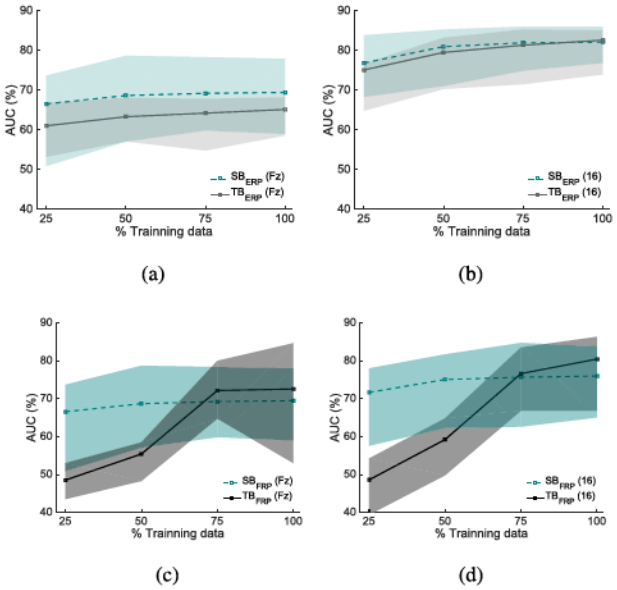

Fig. 5. Expected AUC values for 10 users calculated using different signal models for different calibration lengths ({25 %, 50 %, 75%, and 100%} to train: (a) the single channel (Fz) ERP sequences model, (b) 16 channels ERP sequences model, (c) the single channel (Fz) FRP sequences model, and (d) 16 channels FRP sequences model. The shaded area represents 80% confidence of the cdf of the AUC values for 10 users.

individually. The SB signal model used in the experiments has 25 parameters per EEG channel including the AR parameters, $\mathbf{a}_K$ for $K = 8$, evoked potential waves amplitudes, $\boldsymbol{\beta}$, σ_ξ^2, and hyperparamters. In the current naive approach to multi-channel fusion, the number of parameters grows linearly with the number of channels. For the TB signal model, however, the number of parameters is $N_{P_{TB}} = (d(d-1)/2)+d$, where d is the dimensionality of the EEG feature vector. For a single channel, typically d is approximately 40, and it grows linearly with the number of channels. The multi-channel TB model used in the experiments has approximately 280,000 parameters. The impact of significantly reduced model complexity is evident in results discussed in section IV-D. The SB signal model achieved better than or similar performance to the TB signal model when trained with a smaller number of samples.

V. Conclusions

The performance of EEG-based BCIs relies strongly on the quality of EEG models that are used in human intent detection. Specifically, for designs that employ evoked response potentials, existing approaches typically ignore the temporal dependency of EEG features across trials in a sequence. The results we presented in this paper indicate clearly that ignoring the temporal dependency of EEG features, especially in cases where rapid stimulus sequences are used, may have a huge detrimental effect on performance as measured by information transfer rate. An EEG signal model that attempts to capture the temporal statistical behavior of EEG, as in the case of the here-proposed model, can achieve significant improvements in performance. Furthermore, results indicate that such parametric models can achieve better performance when trained with fewer calibration samples, thus calibration times may be reduced.

References

[1] L. A. Farwell and E. Donchin, "Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalograph. Clin. Neurophysiol.*, vol. 70, no. 6, pp. 510–523, Dec. 1988.

[2] M. Akcakaya *et al.*, "Noninvasive brain-computer interfaces for augmentative and alternative communication," *IEEE Rev. Biomed. Eng.*, vol. 7, pp. 31–49, 2014.

[3] M. Moghadamfalahi, U. Orhan, M. Akcakaya, H. Nezamfar, M. Fried-Oken, and D. Erdoğmuş, "Language-model assisted brain computer interface for typing: A comparison of matrix and rapid serial visual presentation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 23, no. 5, pp. 910–920, Sep. 2015.

[4] L. Acqualagna, M. S. Treder, M. Schreuder, and B. Blankertz, "A novel brain-computer interface based on the rapid serial visual presentation paradigm," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol.*, Sep. 2010, pp. 2686–2689.

[5] U. Orhan, K. E. Hild, D. Erdoğmuş, B. Roark, B. Oken, and M. Fried-Oken, "RSVP keyboard: An EEG based typing interface," in *Proc. Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2012, pp. 645–648.

[6] J. Chmura, J. Rosing, S. Collazos, and S. J. Goodwin, "Classification of movement and inhibition using a hybrid BCI," *Frontiers Neurorobotics*, vol. 11, p. 38, Aug. 2017.

[7] B. J. Edelman, J. Meng, N. Gulachek, C. C. Cline, and B. He, "Exploring cognitive flexibility with a noninvasive BCI using simultaneous steady-state visual evoked potentials and sensorimotor rhythms," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 5, pp. 936–947, May 2018.

[8] T. Zeyl, E. Yin, M. Keightley, and T. Chau, "Improving bit rate in an auditory BCI: Exploiting error-related potentials," *Brain-Comput. Inter.*, vol. 3, no. 2, pp. 75–87, 2016.

[9] R. Yousefi, A. R. Sereshkeh, and T. Chau, "Exploiting error-related potentials in cognitive task based BCI," *Biomed. Phys. Eng. Express*, vol. 5, no. 1, 2018, Art. no. 015023.

[10] A. Cruz, G. Pires, and U. J. Nunes, "Double ErrP detection for automatic error correction in an ERP-based BCI speller," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 1, pp. 26–36, Jan. 2018.

[11] R. Chavarriaga, A. Sobolewski, and J. D. R. Millán, "Errare machinale est: The use of error-related potentials in brain-machine interfaces," *Frontiers Neurosci.*, vol. 22, no. 8, p. 208, Jul. 2014.

[12] N. M. Schmidt, B. Blankertz, and M. S. Treder, "Online detection of error-related potentials boosts the performance of mental typewriters," *BMC Neurosci.*, vol. 13, no. 1, p. 19, 2012.

[13] B. Dal Seno, M. Matteucci, and L. Mainardi, "Online detection of p300 and error potentials in a BCI speller," *Comput. Intell. Neurosci.*, vol. 2010, p. 11, Jan. 2010.

[14] M. Spuler, W. Rosenstiel, and M. Bogdan, "Online adaptation of a c-VEP brain-computer interface (BCI) based on error-related potentials and unsupervised learning," *PLoS One*, vol. 7, no. 12, 2012, Art. no. e51077.

[15] P. Gonzalez-Navarro, M. Moghadamfalahi, M. Akcakaya, and D. Erdoğmuş, "Error-related potentials for EEG-based typing systems," in *Proc. 6th Int. Brain-Comput. Interface Meeting*, 2016.

[16] A. Koçanaoğullari, Y. M. Marghi, M. Akçakaya, and D. Erdoğmuş, "Optimal query selection using multi-armed bandits," *IEEE Signal Process. Lett.*, vol. 25, no. 12, pp. 1870–1874, Dec. 2018.

[17] R. A. Davis, *Gaussian Process*. Hoboken, NJ, USA: Wiley, 2006. [Online]. Available: http://dx.doi.org/10.1002/9780470057339.vag002

[18] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. Cambridge, MA, USA: MIT Press, 2005.

[19] Y. M. Marghi, P. Gonzalez-Navarro, B. Azari, and D. Erdoğmuş, "A parametric EEG signal model for BCIs with rapid-trial sequences," in *Proc. 40th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2018, pp. 118–122.

[20] S. Sur and V. K. Sinha, "Event-related potential: An overview," *Ind. Psychiatry J.*, vol. 18, no. 1, pp. 70–73, 2009.

[21] G. F. Woodman, "A brief introduction to the use of event-related potentials in studies of perception and attention," *Attention, Perception, Psychophys.*, vol. 72, no. 8, pp. 2031–2046, 2010.

[22] U. Orhan *et al.*, "Probabilistic simulation framework for EEG-based BCI design," *Brain-Comput. Inter.*, vol. 3, no. 4, pp. 171–185, 2016.

[23] P. Gonzalez-Navarro, M. Moghadamfalahi, M. Akcakaya, and D. Erdoğmuş, "Spatio-temporal EEG models for brain interfaces," *Signal Process.*, vol. 131, pp. 333–343, Feb. 2017.