

Estimation and Modeling of EEG Amplitude-Temporal Characteristics Using a Marked Point Process Approach

Carlos A. Loza, Jose C. Principe

Abstract—We propose a novel interpretation of single channel Electroencephalogram (EEG) traces based on the transient nature of encoded processes in the brain. In particular, the proposed framework models EEG as the output of the noisy addition of temporal, reoccurring, transient patterns known as phasic events. This is not only neurophysiologically sound, but it also provides additional information that classical EEG analysis often disregards. Furthermore, by utilizing sparse decomposition techniques, it is possible to obtain amplitude and timing that is further modeled using estimation and fitting techniques. We model Brain-Computer Interfaces (BCI) competition data features as Gaussian Mixture Model (GMM) samples in order to show the potential of working in the joint space of the parameters. The results not only preserve the topographic discriminant behavior but also expand the realm of possible EEG analysis.

Index Terms—EEG, Gaussian Mixture Model, Phasic Events, Transient Model

I. INTRODUCTION

EEG characterizes the synchronous average activity of cortical neural populations in the brain. It is the result of the spatiotemporal interactions between graded excitatory postsynaptic potentials (EPSP's) and inhibitory postsynaptic potentials (IPSP's) [1]; thus, it portrays the macroscopic neuronal activity at a relatively lower temporal scale than spikes. In this way, it is complimentary to single neuron models by providing not only amplitude information that is absent in action potentials, but also phase and modulation-related activity that reflects the local interactions between cortical principal cells and interneurons [2].

Due to its noninvasive nature, high temporal resolution and multichannel recording paradigm, EEG has been widely utilized in several studies and applications where it is possible to decode and interpret neural activity. Relevant examples include diagnosis of pathological conditions, such as epilepsy and Parkinson's disease, assessment of sleep stages, and BCI [3]–[6].

Most of these applications rely on the heavy assumptions of stationarity and ergodicity. This is mostly convenient from a statistical point of view because expected values can be replaced by temporal averages in the hope to obtain consistent estimators. However from a neurophysiological stance, it contradicts some of the most important facts regarding neuronal activity. For instance, electrical potentials from the brain are a well-known example of non-stationary stochastic

random processes where the statistical properties vary over time in a transient scheme [7] as a direct consequence of the ongoing reorganization of neuronal assemblies and modulation of their level of synchronism. It is imperative, then, to incorporate this biological constraint into current EEG processing frameworks.

Brockmeier and Principe proposed a model for neuronal oscillations based on the idea that a single EEG trace is the result of transient, reoccurring patterns over time added on a noisy background [8], [9]. The model, since then, has been reformulated in order to incorporate the EEG physiological rhythms; this facilitates the interpretation and opens the possibility for cross-rhythm analysis. Moreover, the current work focuses on the detection of relevant encoded behavior-related patterns over time, also known as phasic events; this task is accomplished without appealing to window-based methods that blur the temporal information and, commonly, provide amplitude data alone. Also, estimation techniques are used to fit the newly discovered parameters in order to quantify and compare them topographically and physiologically.

The rest of the paper is organized as follows: Section 2 describes the transient model for EEG while Section 3 explains the required algorithms needed to determine relevant phasic events. Section 4 details the estimation procedure and illustrates the results using BCI competition data. Lastly, Section 5 concludes the paper and discusses further research.

II. TRANSIENT MODEL FOR EEG

Following the clinical interpretation regarding EEG [7], an anthropomorphic model for neuronal oscillations is established. In particular, a single EEG trace is posed as the result of the noisy addition of reoccurring, transient events over time. From a systems point of view, the transient events can be modeled as shifted versions of finite impulse response (FIR) bandpass filters that will be activated via samples from a pulse train. Moreover, this pulse train is further multiplexed to accommodate the well-known EEG rhythms [10]; in this way, the FIR filters are naturally grouped into filter banks with similar central frequencies.

Fig. 1 depicts the block diagram of the aforementioned model. It is worth noting that each pulse not only contains superior temporal information (up to the inverse of the sampling rate), but it also incorporates amplitude and indexing features assuming each bandpass filter is normalized. Hence, the pulse train can be modeled as a collection of samples from a Marked Point Process (MPP). In equations, the output of the system, $x(t)$, will be an EEG-like signal that resembles the changing structure of neuronal oscillations.

Carlos A. Loza and Jose C. Principe are with the Computational NeuroEngineering Laboratory (CNEL), University of Florida, Gainesville, FL (email: cloza@ufl.edu, principe@cnel.ufl.edu).

This work was supported by Michael J. Fox Grant 9558

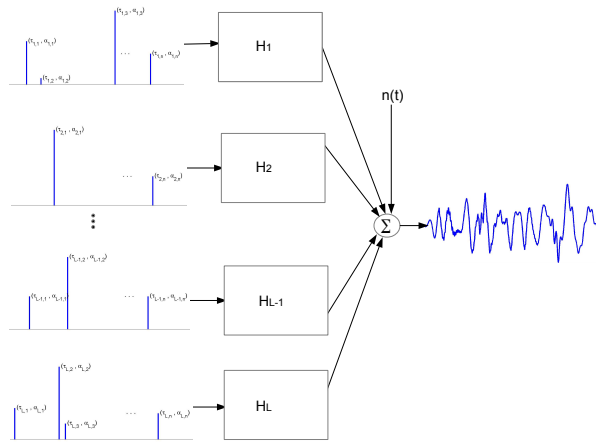


Fig. 1: Transient Model for EEG. A single-channel EEG recording can be regarded as the noisy addition of transient, finite, reoccurring events over time and different frequency bands.

$$x(t) = n(t) + \hat{x}(t) = n(t) + \sum_{i=1}^F y_i(t) \quad (1)$$

$$y_i(t) = \sum_{j=1}^{n_i} \int_{-\infty}^{\infty} \alpha_{i,j} \delta(t - \tau_{i,j}) h_{i,\omega_j} du \quad (2)$$

where F is the number of filters banks, i.e. EEG rhythms to be analyzed, $n(t)$ is the additive noise, in this case colored noise due to the low-pass filtering property of cortical tissue in the brain [11]. α and τ constitute the phasic event's amplitudes and timings, respectively. Lastly, $H_i = \{h_{i,\omega_j}\}$ is the i -th filter bank with unit-norm elements, also known as atoms or patterns. In addition, the collection of MPP samples will also be referred to as the sources of the system.

The paramount advantage of this model is its superior temporal resolution that opens the door for precise detection of phasic events, assessment of cross-rhythm interaction, and computation of relevant statistics beyond the almost pervasive Power Spectral Density (PSD) estimators that rely on power-based measures alone and disregard any information encoded in the phase of the signal. In addition, our transient model can incorporate additional features inherent to the filter banks themselves, e.g. modulation measures, central frequency, and duration. This makes the transient model not only neurophysiologically sound, but also, richer in terms of the potential features obtained for further analysis.

III. DETECTION AND DISCOVERY OF PHASIC EVENTS

Once the model is defined, two different pathways can be adopted: synthesis or analysis. The first one implies the generation of EEG-like time series as a result of discrete samples from the ideal joint distribution of amplitude, timing and indexing characteristics along with the EEG rhythm-based filter banks, and the noise model; for an example of the synthesis scheme, see [12]. On the other hand, the analysis pathway strives to estimate all the model parameters from

a single-channel EEG trace. Specifically, the objective is to find the precise timing, amplitude and additional features of the relevant EEG phasic events.

Different approaches can be adopted to estimate the model parameters. For instance, the simplest solution involves parametric and non-parametric PSD estimation techniques [13]; however, the end result completely disregards the temporal structure of the data due to the one-to-many mapping between the estimated PSD magnitude and the input time series. Alternatively, Time-Frequency (TF) analyses attempt to solve this limitation; nevertheless, the introduction of a processing window smears the temporal information according to the well known time-frequency resolution trade-off. In addition, Short-Time Fourier Transform (STFT) methods and Wavelet Transforms [14] impose complex sinusoids and wavelets as filter banks, respectively. This, in consequence, limits the pattern's shape and derives in overrepresented phasic events.

A more viable solution comes from the inherent sparsity of the MPP; this leads to a reformulation of the analysis problem using a sparse approximation framework [15], i.e. for vectors in a Hilbert space C^Λ :

$$\min_{\Lambda \subset \Omega} \min_{\mathbf{b} \in C^\Lambda} \left\| \mathbf{x} - \sum_{\lambda \in \Lambda} \mathbf{b}_\lambda \varphi_\lambda \right\|_2 \text{ subject to } |\Lambda| \leq L \quad (3)$$

where \mathbf{x} is the input vector, $D = \{\varphi_\omega : \omega \in \Omega\}$ is the dictionary matrix with columns known as atoms; lastly, \mathbf{b} is a list of complex-valued coefficients. The constrained optimization is combinatorial, NP-hard; hence, a computationally fast, greedy algorithm is commonly utilized to obtain a local solution of (3). This technique is known as Matching Pursuit (MP) [16] and is detailed in Algorithm 1 for the time series case where FFT-based operations are computed to assess the locally maximum correlated atom in the dictionary along with its amplitude and timing for each iteration.

MP has been successfully applied to EEG recordings in the past [17]; however, the final representation will heavily rely on the dictionary atoms and the parameter L , i.e. if it is chosen too small, relevant phasic events might be ignored, while if it is chosen too large, the phasic events become overrepresented and the sparsity constraint becomes meaningless. For a possible solution of this problem using alternative sparsity measures, see [12].

Algorithm 1 Matching Pursuit, Time Series Case

Parameters: L

```

 $r(t) \leftarrow x(t)$ 
for  $i = 1 \dots L$  do
   $b_q(t) = \text{xcorr}(\varphi_q(t), r(t)) \quad q = 1, \dots, K$ 
   $p_i \leftarrow \text{argmax}_q \max_t |b_q(t)|$ 
   $\tau_i \leftarrow \text{argmax}_t |b_{p_i}(t)|$ 
   $\alpha_i \leftarrow b_{p_i}(\tau_i)$ 
   $r(t) \leftarrow r(t) - \int_{-\infty}^{\infty} \alpha_i \delta(t - \tau_i - u) \varphi_{p_i}(u) du$ 
end for

```

Algorithm 2 Phasic Event Estimation

Parameters: M, δ $L \leftarrow \lceil N/M \rceil$ $D \leftarrow \text{Dinit}(x(t), 20 \times L, M)$ $D \leftarrow \text{Drest}(D)$ **repeat** $\{(p_i, \alpha_i, \tau_i)\}_{i=1}^L \leftarrow \text{MP}(x(t), D, L)$ $e(t) \leftarrow x(t) - \text{rebuild}(\{(\alpha_i, \tau_i, d_{p_i})\}_{i=1}^L)$ **for** $k = 1, \dots, K$ **do** $w_k \leftarrow \{i | 1 \leq i \leq N, \{p_i\}_{i=1}^L = k\}$ $\mathbf{e}^{(k)} \leftarrow e(t) + \text{rebuild}(\{(\alpha_j, \tau_j, d_{p_j})\}_{j \in w_k}^L)$ $E_k \leftarrow \text{matrix}(\mathbf{e}^{(k)}, \{\tau_j\}_{j \in w_k}, M)$ $E_k \leftarrow U\Delta V^T$ $d_{p_i} \leftarrow U(:, 1)$ **end for****until** convergence $\theta \leftarrow 0, J \leftarrow 1$ **while** $\theta < \delta$ **do** $\theta \leftarrow \text{norm}(x(t)) / \text{norm}(\text{rebuild}(\{(\alpha_i, \tau_i, d_{p_i})\}_{i=1}^L))$ $J \leftarrow J + 1$ **end while** $L \leftarrow J - 1$ $\{(p_i, \alpha_i, \tau_i)\}_{i=1}^L \leftarrow \text{MP}(x(t), D, L)$

It would be more advantageous and principled to detect the relevant phasic events from the EEG itself without depending on predefined filter banks; this builds a generative dictionary from the sparse patterns over time. A possible solution involves dictionary learning techniques, such as K-SVD [18] where alternating optimizations are performed: first, MP sparse approximation, and then, dictionary learning using Singular Value Decomposition (SVD).

Brockmeier applied this technique to neuronal oscillations in the past [8]; however, we improve the phasic event discovery by limiting the potential candidates for dictionary components. Specifically, the function *Dinit* computes a large number ($20 \times L$ as heuristic) of atoms based on shifted M -sample long snippets from the input time series. Here, L indicates the maximum possible number of non-overlapping phasic events present in the EEG trace. Moreover, based on clinical interpretations [19] and the temporal interactions of firings between pyramidal cells and interneurons [1], we are mostly interested in spindle-shaped phasic events that reflect amplitude modulation over time. This is achieved by the function *Drest* that computes a smooth, interpolated version of the Hilbert Transform magnitude [20] of the atoms and discards the filters with non-spindle-like shape. This not only constrains the modulation pattern of the relevant phasic events, but also accelerates the overall performance of the algorithm by reducing the size of the dictionary.

As algorithm 2 details, the alternating optimization problem first fixes the dictionary and performs sparse approximation; then, a matrix of patches, E_k , with only the active k -th atoms is built. Afterward, that particular k -th pattern is updated via SVD of E_k . The process is repeated for

each atom in the dictionary. The alternating optimization is stopped after convergence is reached in the form of a running variance threshold (1×10^{-5}). Lastly, several reconstructed signals are computed via the function *rebuild* until the normalized reconstructed signal power threshold δ is reached.

Fig. 2 shows an example of the final estimated phasic events. Specifically, the data corresponds to BCI competition dataset IIIb, subject O3VR [21]. For this experiment, the user is asked to imagine a left/right arm movement in order to control a cursor in a screen in front of him. Thus, two different labels are provided along with two bipolar recordings corresponding to C3 and C4. It is worth mentioning that throughout this paper, all the analysis and processing was performed utilizing a bandpass filtered version of the provided traces. We focus on the μ rhythm that is known for motor-related activity encoding [22]. Lastly, $M = 50$ samples (0.4 s.) is chosen as the duration of the phasic events.

IV. ESTIMATION OF MODEL PARAMETERS

Once the phasic events are isolated, it is necessary to model the MPP parameters. We focus on the 2-second sequence in [21] involving a trigger signal at $t = 0$ that acts as a pre-stimulus and a visual cue at $t = 1$ when the subject is instructed to perform the imagery movement task. The first results compare the log-variance computed from the original bandpass filtered EEG and the log-decomposition α 's. The logarithm transform was computed to encourage Gaussianity and for the phasic event discovery algorithm, we heuristically set $\delta = 0.9$. To summarize, the average values along with the significance relevance (t-test) are shown in Table I. As the p-values suggest, the MP, K-SVD decomposition is able to preserve the statistical significance, and thus discriminability, of the magnitude between channels across conditions.

TABLE I: Average log-variance and log- α per task and channel

	Task 1		Task 2	
	C3	C4	C3	C4
Log-var	3.95	1.98	2.18	3.94
p-value	1.89×10^{-11}		7.95×10^{-9}	
Log-α	10.70	9.61	9.71	10.69
p-value	2.68×10^{-8}		3.48×10^{-7}	

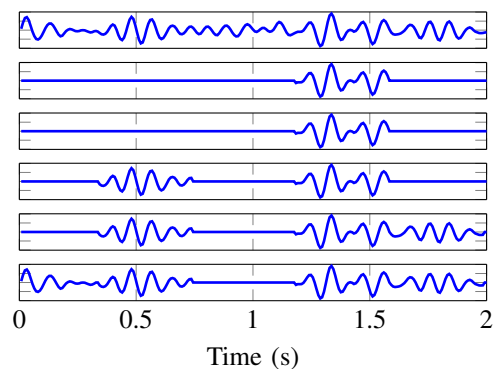


Fig. 2: Example of phasic event discovery. From top to bottom: Bandpass filtered EEG trace. 1 phasic event detected, $\delta = 0.5$. 1 phasic event detected, $\delta = 0.6$. 2 phasic events detected, $\delta = 0.7$. 3 phasic events detected, $\delta = 0.8$. 4 phasic events detected, $\delta = 0.9$.

One important fact should be noted at this point: the 2-second interval under analysis includes two very distinct behavioral stages, i.e. readiness or pre-movement and imagery movement itself. However, Table I does not reflect this phase transition due to the absence of temporal information. Hence, we explore the joint (α, τ) distribution obtained from algorithm 2 and fit it using the Expectation Maximization (EM) algorithm [23] under Gaussian priors in order to elucidate this matter. Additionally, the optimal number of components is assessed using the Bayesian Information Criterion (BIC) [24]; in this way, the 2-dimensional pdf is posed as a GMM with P components and weights $\{q_i\}_{i=1}^P$.

For instance, Fig. 3 depicts the histograms and scatter plot of the aforementioned distributions for electrode C3 and task 1. In addition, Table II describes the 3-component multivariate GMM parameters and significance for Task 2.

TABLE II: Topographic GMM parameters ($\bar{\mu}_q = [\mu_{\tau_q}, \mu_{\alpha_q}]$) for right hand movement (Task 2).

	q_1	$\bar{\mu}_1$	q_2	$\bar{\mu}_2$	q_3	$\bar{\mu}_3$
C3	0.14	[0.2, 9.8]	0.69	[1.0, 9.6]	0.17	[1.8, 9.9]
C4	0.17	[0.2, 10.0]	0.66	[1.0, 10.8]	0.17	[1.7, 10.9]
p-value	4.62×10^{-1}		8.74×10^{-17}		1.94×10^{-3}	

Statistical relevance is computed using the two-sample Hotelling's T-square test for each component. In this way, it is evident that the discriminant behavior is not significant at the beginning of the trial and reaches its apex around the 1-second mark; this is, in fact, expected according to the experiment setting. Task 1 displays similar properties but is omitted for brevity.

V. CONCLUSIONS AND FURTHER WORK

We have modeled the joint pdf of amplitudes and timings regarding μ rhythm phasic events. This opens the possibility for novel EEG analysis with exceptional temporal resolution. For instance, the relationship between microscopic activity, such as spikes, and macroscopic phenomena or neuronal

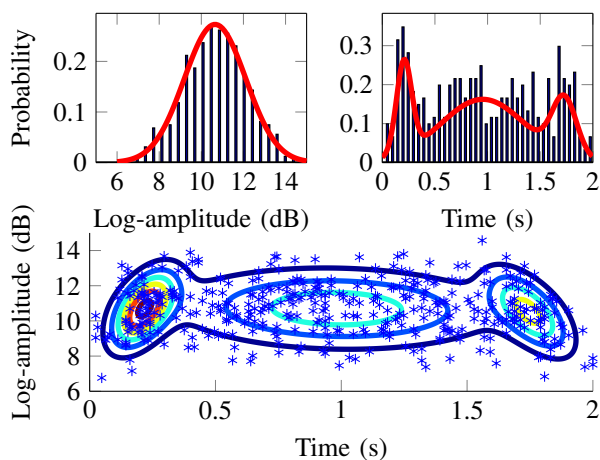


Fig. 3: Example of decomposition models. Top left: $\log\alpha$ with estimated Gaussian model. Top right: τ with estimated GMM ($P = 3$). Bottom: $\log\alpha, \tau$ with estimated GMM ($P = 3$)

oscillations can be directly assessed using temporal point process techniques. In addition, the joint space can be utilized to compute distance metrics in order to improve discriminability and quantify differences across topographical regions of the brain under diverse behavioral states.

REFERENCES

- [1] E. Niedermeyer and F. L. da Silva, *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins, 2005.
- [2] G. Buzsáki, C. A. Anastassiou, and C. Koch, "The origin of extracellular fields and currents EEG, ECoG, LFP and spikes," *Nature reviews neuroscience*, vol. 13, no. 6, pp. 407–420, 2012.
- [3] S. Smith, "EEG in the diagnosis, classification, and management of patients with epilepsy," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 76, no. suppl 2, pp. ii2–ii7, 2005.
- [4] R. Soikkeli, J. Partanen, H. Soininen, A. Pääkkönen, and P. Riekkinen, "Slowing of EEG in parkinson's disease," *Electroencephalography and clinical neurophysiology*, vol. 79, no. 3, pp. 159–165, 1991.
- [5] A. Rechtschaffen and A. Kales, "A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects," 1968.
- [6] M. A. Lebedev and M. A. Nicolelis, "Brain-machine interfaces: past, present and future," *TRENDS in Neurosciences*, vol. 29, no. 9, pp. 536–546, 2006.
- [7] G. Buzsáki and A. Draguhn, "Neuronal oscillations in cortical networks," *science*, vol. 304, no. 5679, pp. 1926–1929, 2004.
- [8] A. J. Brockmeier, "Learning and exploiting recurrent patterns in neural data," Ph.D. dissertation, Citeseer, 2014.
- [9] J. C. Principe and A. J. Brockmeier, "Representing and decomposing neural potential signals," *Current opinion in neurobiology*, vol. 31, pp. 13–17, 2015.
- [10] G. Buzsáki, *Rhythms of the Brain*. Oxford University Press, 2006.
- [11] H. Lindén, K. H. Pettersen, and G. T. Einevoll, "Intrinsic dendritic filtering gives low-pass power spectra of local field potentials," *Journal of computational neuroscience*, vol. 29, no. 3, pp. 423–444, 2010.
- [12] C. A. Loza and J. C. Principe, "Transient model of EEG using gini index-based matching pursuit," in *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*. IEEE, 2016, p. to appear.
- [13] P. Stoica and R. L. Moses, *Spectral analysis of signals*. Pearson/Prentice Hall Upper Saddle River, NJ, 2005.
- [14] L. Cohen, *Time-frequency analysis*. Prentice hall, 1995, vol. 299.
- [15] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *Information Theory, IEEE Transactions on*, vol. 50, no. 10, pp. 2231–2242, 2004.
- [16] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *Signal Processing, IEEE Transactions on*, vol. 41, no. 12, pp. 3397–3415, 1993.
- [17] P. Durka and K. Blinowska, "Analysis of eeg transients by means of matching pursuit," *Annals of biomedical engineering*, vol. 23, no. 5, pp. 608–611, 1995.
- [18] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *Signal Processing, IEEE Transactions on*, vol. 54, no. 11, pp. 4311–4322, 2006.
- [19] W. O. Tatum IV, *Handbook of EEG interpretation*. Demos Medical Publishing, 2014.
- [20] S. L. Hahn, *Hilbert transforms in signal processing*. Artech House on Demand, 1996.
- [21] B. Blankertz, K.-R. Müller, D. J. Krusienski, G. Schalk, J. R. Wolpaw, A. Schlögl, G. Pfurtscheller, J. R. Millan, M. Schröder, and N. Birbaumer, "The BCI competition iii: Validating alternative approaches to actual BCI problems," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 14, no. 2, pp. 153–159, 2006.
- [22] G. Pfurtscheller, C. Brunner, A. Schlögl, and F. L. Da Silva, "Mu rhythm (de) synchronization and EEG single-trial classification of different motor imagery tasks," *NeuroImage*, vol. 31, no. 1, pp. 153–159, 2006.
- [23] T. K. Moon, "The expectation-maximization algorithm," *Signal processing magazine, IEEE*, vol. 13, no. 6, pp. 47–60, 1996.
- [24] G. Schwarz *et al.*, "Estimating the dimension of a model," *The annals of statistics*, vol. 6, no. 2, pp. 461–464, 1978.