COPING WITH TRIAL-TO-TRIAL VARIABILITY OF EVENT RELATED SIGNALS: A BAYESIAN INFERENCE APPROACH

Mingzhou Ding[a], Kevin Knuth[b], Yonghong Chen[a], Steven L. Bressler[c], and Charles E. Schroeder[d]

a Department of Biomedical Engineering, University of Florida, Gainesville, FL 32611, USA;
b NASA Ames Research Center, Moffett Field, CA 94035, USA;
c Center for Complex Systems and Brain Sciences, Florida Atlantic University, Boca Raton, FL 33431, USA;
d Nathan Kline Institute, Orangeburg, NY 10962, USA

ABSTRACT

In electroneurophysiology, single-trial brain responses to a sensory stimulus or a motor act are commonly assumed to result from the linear superposition of a stereotypic event-related signal (e.g. the event-related potential or ERP) that is invariant across trials and some ongoing brain activity often referred to as noise. To extract the signal, one performs an ensemble average of the brain responses over many identical trials to attenuate the noise. To date, this simple signal-plus-noise (SPN) model has been the dominant approach in cognitive neuroscience. Mounting empirical evidence has shown that the assumptions underlying this model may be overly simplistic. More realistic models have been proposed that account for the trial-to-trial variability of the event-related signal as well as the possibility of multiple differentially varying components within a given ERP waveform. The variable-signal-plus-noise (VSPN) model, which has been demonstrated to provide the foundation for separation and characterization of multiple differentially varying components, has the potential to provide a rich source of information for questions related to neural functions that complement the SPN model. Thus, being able to estimate the amplitude and latency of each ERP component on a trial-by-trial basis provides a critical link between the perceived benefits of the VSPN model and its many concrete applications. In this paper we describe a Bayesian approach to deal with this issue and the resulting strategy is referred to as the differentially Variable Component Analysis (dVCA). We compare the performance of dVCA on simulated data with Independent Component Analysis (ICA) and analyze neurobiological recordings from monkeys performing cognitive tasks.

1. INTRODUCTION

Relations between brain and behavior are often studied by recording event-related potentials (ERPs) over repeated presentations of a sensory stimulus or task performance. Each task performance is referred to as a trial. On a single-trial basis, the traditional model holds that the data from a given electrode (channel) is the linear superposition of an event-related potential (ERP), which is considered the signal having a characteristic waveform whose amplitude and latency stay the same each time the event is repeated on multiple trials, and ongoing background activity (noise). We refer to this model as the signal plus noise (SPN) model and represent it mathematically as [1].

$$x_r(t) = s(t) + \eta_r(t) \tag{1}$$

This work supported by NIMH grants MH070498, MH71620 and MH64204, and the NASA SISM Intelligent Systems Program.

where $x_r(t)$ is the EEG recording from the r^{th} trial, $s(t)$ is the ERP, and $\eta_r(t)$ is the ongoing noise process. The empirical evidence suggests that the SPN model is overly simplistic and a more realistic model should capture the trial-to-trial variability in amplitude and latency of the event-related signals and the existence of multiple components with differential variability in their single-trial amplitudes and latencies. A model that incorporates all these features, which we call the variable signal plus noise (VSPN) model, can be written as [2]

$$x_r(t) = \sum_{n=1}^{N} \alpha_{nr} s_n(t - \tau_{nr}) + \eta_r(t) \tag{2}$$

where $s_n(t)$ is the n^{th} event-related component waveform with trial-to-trial variable amplitude and latency given by α_{nr} and τ_{nr}, respectively, and N is the total number of components. In many experimental paradigms, investigators use multiple detectors positioned at different locations to measure the event-related activities of neural ensembles. In this case, there is a need to consider the specific neural ensemble whose activity is responsible for a given component that appears across all detectors. We call such an ensemble the source or the generator for the component. The degree to which a detector records a signal from a particular source depends on many factors including the relative position and orientation of the source relative to the detector. To describe this source-detector interaction, we introduce a coupling matrix $\mathbf{C}$, where the matrix element C_{mn} describes the degree to which the m^{th} detector $(m = 1, 2, \ldots, M)$ detects the n^{th} source. This coupling matrix is known as the mixing matrix in the source separation literature and as the lead-field matrix in electroneurophysiology. For the r^{th} recorded trial we model the signal $x(t)$ recorded by the m^{th} detector in component form as

$$x_{mr}(t) = \sum_{n=1}^{N} C_{mn} \alpha_{nr} s_n(t - \tau_{nr}) + \eta_{mr}(t) \tag{3}$$

where the notation has the same interpretation as before (see Equation (2)). For the purpose of this work we will refer to both the single sensor model in (2) and multi-sensor model in (3) as the VSPN model. The goal of this work is to present a Bayesian inference framework for the estimation of all the parameters in the VSPN models and to then use the trial-to-trial variability information of the ERP components to gain understanding of neural functions.

2. METHODS

For simplicity we consider in more detail the single channel VSPN model in (2). Bayesian inference for the parameters in the VSPN model in (3) can be similarly formulated. The posterior probability density function (PPDF) is formulated as

$$p(\mathbf{s},\alpha,\tau,\theta_\eta|\mathbf{x},I) \propto p(\mathbf{s},\alpha,\tau,\theta_\eta(t)|I)\times \frac{p(\mathbf{x}|\mathbf{s},\alpha,\tau,\theta_\eta(t),I)}{p(\mathbf{x}|I)} \quad (4)$$

where the boldface parameters denote the set of parameters for all the components over the whole ensemble of trials. The parameter $\theta_\eta(t)$ denotes the parameters for the ongoing process and $p(\mathbf{s},\alpha,\tau,\theta_\eta(t)|I)$ is the prior probability for the model parameters. For this additive model, the likelihood $p(\mathbf{x}|\mathbf{s},\alpha,\tau,\theta_\eta(t),I)$ is given by the probability model of the ongoing activity, i.e., $p(\eta(t)|I)$. In the absence of precise knowledge about the temporal structure of the ongoing activity, we assign $\eta(t)$ to be independent identically distributed with a (unknown) time-independent variance σ_η^2 and zero mean. In this way, Equation (4) is rewritten as

$$p(\mathbf{s},\alpha,\tau,\sigma_\eta|\mathbf{x},I) \propto p(\mathbf{s},\alpha,\tau,\sigma_\eta|I)\frac{p(\eta|\sigma_\eta,I)}{p(\mathbf{x}|I)}$$

Under the constraint of a given mean and σ_η^2, and following the principle of maximum entropy [3], a Gaussian density is assigned to the likelihood function. After dropping the normalization term $p(\mathbf{x}|I)^{-1}$, the PPDF can be rewritten as

$$p(\mathbf{s},\alpha,\tau,\sigma_\eta|\mathbf{x},I) \propto p(\mathbf{s},\alpha,\tau,\sigma_\eta|I)\times\left(2\pi\sigma_\eta^2\right)^{-\frac{RT}{2}}\times \exp\left(\frac{-\sum_{r=1}^{R}\sum_{t=1}^{T}\left[x_r(t)-\sum_{n=1}^{N}\alpha_{nr}s_n(t-\tau_{nr})\right]^2}{2\sigma_\eta^2}\right) \quad (5)$$

where R is the total number of trials, T is the total number of sampled data points in a given trial and other parameters have the same interpretation as before. For simplicity we take the prior distributions of α_{nr}, τ_{nr}, and $s_n(t)$ to be uniform, with appropriate cutoffs reflecting physiologically reasonable ranges of values. Treating the variance of the ongoing activity as a nuisance parameter and assigning the Jeffreys prior $p(\sigma_\eta|I)=\sigma_\eta^{-1}$ [3], we marginalize the posterior over σ_η and obtain

$$p(\mathbf{s},\alpha,\tau|\mathbf{x},I) \propto p(\mathbf{s},\alpha,\tau|I)\times \left(2\pi\sum_{r=1}^{R}\sum_{t=1}^{T}\left[x_r(t)-\sum_{n=1}^{N}\alpha_{nr}s_n(t-\tau_{nr})\right]^2\right)^{-\frac{RT}{2}}. \quad (6)$$

A thorough evaluation of the PPDF and computation of its moments can be obtained via Markov Chain Monte Carlo (MCMC) methods [3]. Here we summarize the PPDF by seeking the Maximum a Posteriori (MAP) solution, i.e., a set of parameters that maximize the PPDF. Specifically, the MAP solution for the collection of model parameters M is

$$\hat{M} = \arg\max_M[\ln(p(M|I))+\ln(p(D|M,I))] \quad (7)$$

where D stands for data.

Intuition about the characteristics of the MAP solution can be gained by examining the partial derivatives of the logarithm of the PPDF with respect to each of the model parameters. This leads to a practical and simple estimation algorithm. If we let

$$Q=\sum_{r=1}^{R}\sum_{t=1}^{T}\left[x_r(t)-\sum_{n=1}^{N}\alpha_{nr}s_n(t-\tau_{nr})\right]^2 \quad (8)$$

logarithm of the posterior can be simply written as

$$\ln P = -\frac{RT}{2}\ln Q + const, \quad (9)$$

where P stands for PPDF. Setting partial derivatives against the unknown single-trial parameters to zero we can solve the resulting equations in an algorithmic fashion [2].

The MAP solution above leads to a simple iterative fixed point algorithm. After proper initialization, at each iteration step the parameters for all components are updated in sequence: first the latencies, then the waveforms, and finally the amplitudes. Specifically, let $s_j^i(t)$, α_{jr}^i, τ_{jr}^i denote the estimated values of the parameters of the j^{th} source during the r^{th} trial in the i^{th} iteration. To avoid degeneracy in the model, the averages of the amplitude and latency values in each iteration are constrained to $\langle\alpha_{jr}^i\rangle_r\equiv 1$ and $\langle\tau_{jr}^i\rangle_r\equiv 0$. In this way, if there is no trial-to-trial variability both in amplitude and latency, the superposition of the estimated component waveforms should equal to the simple average of single trial time series referred to here as AERP. For a single channel data set $\mathbf{x}$, the algorithm consists of the following steps:

- At $m=0$, the initial guess for the amplitudes and time offsets are set to $\alpha_{jr}^0=1$, $\tau_{jr}^0=0$. For simplicity, the decision on the number of components N is based on the inspection of the AERP. Similarly, N non-overlapping segments of the AERP are taken as the initial guesses for the N components' waveforms . After this initialization, each iteration consists of four steps:
  - For a given trial, estimate the latency one component at a time, according to $\tau_{jr}^{i+1}=\arg\max\rho^i(\tau)$, with $\rho(t)=\sum_{t=1}^{T}\left[\alpha_{jr}s_j(t-\tau)\left(x_r(t)-\sum_{n=1,n\neq j}^{N}\alpha_{nr}s_n(t-\tau)\right)\right]$ starting from the first and proceeding up to the N^{th} component.
  - Estimate the waveforms according to: $s_j^{i+1}(t)=\frac{\sum_{r=1}^{R}W\alpha_{jr}^i}{\sum_{r=1}^{R}(\alpha_{jr}^i)^2}$, with $W=x_r(t+\tau_{jr}^{i+1})-\sum_{n=1,n\neq j}^{N}\alpha_{jr}^i s_n^i(t-\tau_{nr}^{m+1}+\tau_{jr}^{m+1})$.
  - For all the trials and components, estimate the amplitudes according to: $\alpha_{jr}^{i+1}=\frac{\sum_{t=1}^{T}UV}{\sum_{t=1}^{T}V^2}$, with $U=x_r(t)-\sum_{n=1,n\neq j}^{N}\alpha_{nr}^i s_n^{i+1}(t-\tau_{nr}^{i+1})$ and $V=s_j^{i+1}(t-\tau_{jr}^{i+1})$.
  - Increment the iteration index: $i=i+1$; Repeat 2 through 4 for K iterations until convergence or for a pre-decided number of iterations.

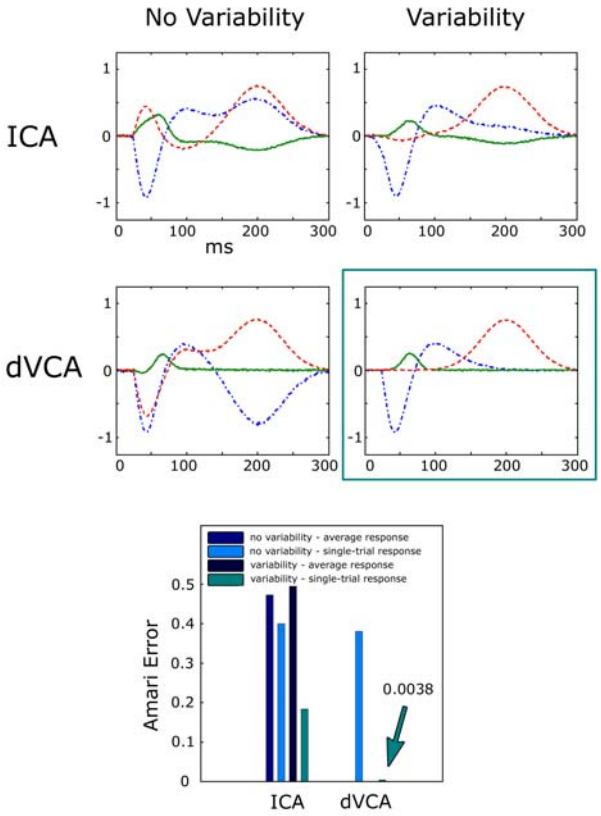

Figure 1: (Top) The source waveshape results for extended ICA, and dVCA in the single-trial analysis cases. Only the dVCA analysis of the variable responses results in accurate identification of the underlying source waveshapes (box). (Bottom) The associated Amari errors for the cases above.

3. COMPARISON TO ICA

The dVCA algorithm shares many similarities to Infomax ICA [4], since in both cases, the sources are assumed to be linearly mixed. The greatest difference is that the signal model used in dVCA is more specialized than the linear mixing model used in ICA; it allows for the sources to vary in specific ways from trial to trial. Second, dVCA solves for all the parameters, whereas ICA marginalizes over the source waveshapes [6, 7, 2]. Third, ICA through the use of the logistic function makes assumptions about the amplitude density of the source waveshapes [5, 6], whereas dVCA makes no such specific assumptions.

The greatest strength of dVCA arises due to the fact that it explicitly models trial-to-trial variability. Here we briefly compare dVCA to ICA, and show that dVCA outperforms ICA when trial-to-trial variability is a significant factor in the data. We simulated electric field potentials recorded from a linear-array multielectrode with 15 channels spanning the cortical laminae in V1 by designing three synthetic ERP components sampled at 2 kHz to approximate the neural ensemble response to diffuse red-light stimulation in macaque V1 (see next section). Two synthetic data sets were used to perform the comparison; both of which were contaminated by low amplitude additive Gaussian noise (component SNRs of $7.2dB$, $-6.7dB$, and $4.5dB$). The first data set exhibited neither amplitude nor latency variability. The second data set exhibited log-normally distributed amplitude variability with sample mean $\mu = 1.0$ and sample standard deviation $\sigma = 1.0$, normally-distributed latency variability with sample mean $\mu = 0.0ms$ and sample standard deviation $\sigma = 10.0ms$.

ICA was performed by applying the Infomax ICA algorithm in the EEGLAB toolbox [8]. The runica.m algorithm was used with the 'extend' option (extended ICA), which allows ICA to model sub-Gaussian as well as super-Gaussian sources. Since each data set has 15 channels, ICA estimates 15 source waveshapes. The three estimated source waveshapes with highest correlation to the three original sources were chosen for analysis while the remaining estimated sources were ignored.

The data sets were analyzed by ICA in two different ways: first by averaging the epochs and analyzing the average response (average analysis), and second by treating the data as a long string of concatenated single-trial responses (single-trial analysis). Figure 1 (top) shows the source waveshape results for extended ICA, and dVCA in the single-trial analysis cases. Only the dVCA analysis of the variable responses results in accurate identification of the underlying source waveshapes. The Amari errors [9] in Figure 1 (bottom) quantify the degree to which the components were correctly identified in the four cases. Since dVCA can only analyze single-trial responses, the dVCA analysis consisted of analyzing the single-trial cases only resulting in only two bars in the bar graph. Note that dVCA performs poorly when there is no trial-to-trial variability. However, the presence of trial-to-trial variability dramatically improves the performance of dVCA enabling it to surpass extended ICA in separation quality.

4. APPLICATION TO EXPERIMENTAL DATA

Two male macaque monkeys were trained to perform an intermodal selective attention task [10]. Streams of interdigitated auditory and visual stimuli were presented with an irregular interstimulus interval (ISI) (minimum of 350 ms). The attended modality was alternated across trial blocks while the physical stimuli remained the same. The eye positions were carefully monitored and controlled within a predefined window for both attending conditions. In both auditory and visual modalities, two types of stimuli were presented: a 'standard' which occurred 86% of the time and a 'deviant' (created by changing the intensity of the standard stimuli), which was presented on 14% of the trials. The visual stimuli were diffusive light flashes subtending a 12-degree visual angle centered at the point of gaze. Attention to a given sensory modality was maintained by requiring the subject to make a lever-release response to the deviant stimuli while ignoring all stimuli in the other modality. The level of performance accuracy was around 92% and difficulty of discrimination was equated across the two modalities. A linear array electrode with 14 contacts (150 micron spacing) spanning all six cortical laminae was inserted in different parts of the visual systems. Laminar LFPs were sampled at 2 kHz. Current source density profiles were computed by taking the second spatial derivative of the laminar LFPs. Here only neural activity from V1 in response to the standard visual stimulus under two different attending conditions will be considered to examine the effect of attention.

A key question in the research on the neural correlates of

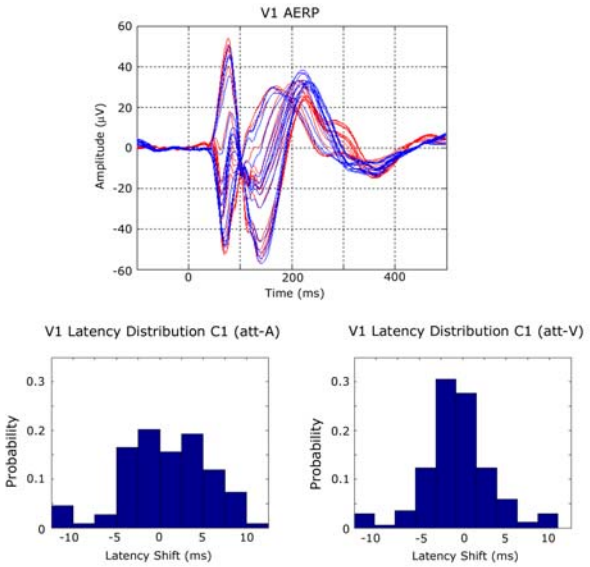

Figure 2: (Top) 14 averaged ERPs along the depth electrode for the attending-visual (att-V) condition (red) and the attending-auditory (att-A) condition (blue). (Bottom) The latency distributions of the responses in the two conditions.

attention is whether early neural responses in primary sensory areas are modulated by attention. Reports on this issue in the primate visual attention literature are conflicted [11, 12], with some reporting attention modulation starting from the response onset while others contending that attentional effect lagging response onset by a substantial time difference [10]. While the main technique previously used is the grand averaging method, we believe that single-trial parameters may provide a greater sensitivity to attentional effects in different experimental paradigms.

In Figure 2 (top) we display the superposition of all 14 averaged ERPs along the depth electrode for the attending-visual condition (red) and the attending-auditory (ignoring-visual) condition (blue). No statistically significant differences are observed for the first component from 50 ms to 100 ms (i.e. early visual processing). To examine whether the single trial parameters are modulated by attention we measured single trial latencies by the dVCA method for the first evoked component. The distributions of the trial-by-trial latencies are shown in Figure 2 (bottom). The distributions for the two conditions are significantly different ($p < 0.01$). In particular, when the monkey attended the visual modality, a narrower and more focused latency distribution is seen (Figure 2 bottom right) when compared to the distribution from attending-auditory trials, suggesting that on a trial-by-trial basis, the timing of the evoked component becomes more reliable. The narrower latency distribution of the first component is often accompanied by smaller trial-by-trial amplitudes (results not shown). On simple averaging this narrow latency distribution is offset by the smaller amplitude distribution and the result is an AERP that is not different from the ignoring-visual condition. The reason why the evoked component has a smaller amplitude for attending-visual trials is not known.

REFERENCES

[1] Truccolo, W.A., Ding, M., Knuth, K.H., Nakamura, R., and Bressler, S.L., "Trial-to-trial variability of cortical evoked responses: implications for the analysis of functional connectivity," *Clin. Neurophysiol.* 113(2), 206-226, 2002

[2] Truccolo, W.A., Knuth, K.H., Shah, A.S., Bressler, S.L., Schroeder, C.E., and Ding, M., "Estimation of single-trial multi-component ERPs: Differentially variable component analysis (dVCA)," *Biol Cybern.* 89, 426-438, 2003

[3] Gelman, A., Carlin, J.B., Stern, H.S., and Rubin, D.B., *Bayesian Data Analysis*, Chapman and Hall, New York, 1995

[4] A. J. Bell and T. J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural Comp.*, vol. 7, pp. 1129–1159, 1995.

[5] K. H. Knuth "Difficulties applying recent blind source separation techniques to EEG and MEG," in G. J. Erickson, J. T. Rychert and C. R. Smith (eds.), *Maximum Entropy and Bayesian Methods, Boise 1997*, Kluwer, Dordrecht, pp. 209-222, 1997.

[6] K. H. Knuth, "A Bayesian approach to source separation," in J.-F. Cardoso, C. Jutten and P. Loubaton (eds.), *Proceedings of the First International Workshop on Independent Component Analysis and Signal Separation: ICA'99*, Aussois, France, Jan. 1999, pp. 283-288, 1999.

[7] K. H. Knuth, W. A. Truccolo, S. L. Bressler and M. Ding "Separation of multiple evoked responses using differential amplitude and latency variability," in T.-W. Lee, T.-P. Jung, S. Makeig, T.J. Sejnowski (eds.), *Proceedings of the Third International Workshop on Independent Component Analysis and Blind Signal Separation (ICA 2001)*, Dec. 9-12, 2001, San Diego CA, pp. 463-468, 2001.

[8] A. Delorme, S. Makeig "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *J. Neurosci. Meth.* 134, 9-21, 2004.

[9] S.-I. Amari, A. Cichocki, H. H. Yang "A new learning algorithm for blind signal separation," in D. Touretzky, M. Mozer, M Hasselmo (eds.), *Advances in Neural Information Processing Systems 8*, Cambridge, MA: MIT Press, pp. 752-763, 1996.

[10] Mehta, A.D., Ulbert, I., and Schroeder, C.E., "Intermodal selective attention in monkeys I: distribution and timing of effects across visual areas," *Cereb. Cortex.* 10(4), 343-58, 2000

[11] Luck, S.J., Chelazzi, L., Hillyard, S.A., and Desimone,R., "Neural mechanisms of spatial selective attention in areas V1, V2 and V4 of macaque visual cortex," *J. Neurophysiol.* 77, 24-42, 1997

[12] Motter, B.C., "Focal attention produces spatially selective processing in visual cortical areas V1, V2 and V4 in the presence of competing stimuli," *J. Neurophysiol.* 70, 909-919,1993