

A Bayesian Approach to Estimating Coupling Between Neural Components: Evaluation of the Multiple Component, Event-Related Potential (mcERP) Algorithm

Ankoor S. Shah^{1,2}, Kevin H. Knuth³, Wilson A. Truccolo⁴, Mingzhou Ding⁵, Steven L. Bressler⁵, Charles E. Schroeder^{1,2}

¹*Department of Neuroscience, Albert Einstein College of Medicine, Bronx, NY 10461 USA*

²*Cognitive Neuroscience and Schizophrenia Program, Nathan Kline Institute, Orangeburg, NY 10962 USA*

³*NASA Ames Research Center, Computational Sciences Department, Code IC Moffett Field, CA 94035 USA*

⁴*Neuroscience Department, Brown University, Providence, RI 04035 USA*

⁵*Center for Complex Systems, Florida Atlantic University, Boca Raton, FL 33431 USA*

Abstract. Accurate measurement of single-trial responses is key to a definitive use of complex electromagnetic and hemodynamic measurements in the investigation of brain dynamics. We developed the multiple component, Event-Related Potential (mcERP) approach to single-trial response estimation to improve our resolution of dynamic interactions between neuronal ensembles located in different layers within a cortical region and/or in different cortical regions. The mcERP model asserts that multiple components defined as stereotypic waveforms comprise the stimulus-evoked response and that these components may vary in amplitude and latency from trial to trial. *Maximum a posteriori* (MAP) solutions for the model are obtained by iterating a set of equations derived from the posterior probability. Our first goal was to use the mcERP algorithm to analyze interactions (specifically latency and amplitude correlation) between responses in different layers within a cortical region. Thus, we evaluated the model by applying the algorithm to synthetic data containing two correlated local components and one independent far-field component. Three cases were considered: the local components were correlated by an interaction in their single-trial amplitudes, by an interaction in their single-trial latencies, or by an interaction in both amplitude and latency. We then analyzed the accuracy with which the algorithm estimated the component waveshapes and the single-trial parameters as a function of these relationships. Extensions of these analyses to real data are discussed as well as ongoing work to incorporate more detailed prior information.

INTRODUCTION

Electromagnetic and hemodynamic measurements of brain activity are recorded to investigate neurophysiological processing of sensory stimuli with the assumption that these signals will yield insight into the brain dynamics of sensation. In the case of electromagnetic measures such as event-related potentials (ERPs), the recorded signal represents the superposition of activity from numerous neural sources on the detector array. Typically there is large trial-to-trial variability in the ERP waveform while the

subject is exposed to repeated presentations or trials of the same stimulus. The recorded ERP signals are then averaged across these trials generating the average ERP. Although response averaging improves the signal-to-noise ratio, it precludes investigation of the dynamic features of the system, including trial-to-trial variability of the evoked response, dynamic interactions between stimulus-evoked components, and stimulus-induced modulation of ongoing electroencephalographic (EEG) rhythms.

The multiple component Event-Related Potential (mcERP) model was introduced as an alternative technique for examining these signals [1]. The mcERP model states that multiple components generate the signal and that each of these components may vary in amplitude and latency from trial to trial. Furthermore, the model makes no assumptions about the interdependency of the different components. Thus, this model estimates the different components generated in response to a stimulus presentation, quantifies single-trial variability in the evoked potential, and permits exploration of dynamic interactions between related components.

One of our goals is to apply the mcERP model to data recorded from the primary visual pathways of the macaque monkey to supplement our understanding of the dynamics of visual processing. Since this is a heavily studied system, the literature provides numerous points on which to index our findings. Anatomical studies outline a “processing hierarchy” based on structured systematic interconnections within and between the brain regions comprising the system [2]. Electrophysiological studies of averaged visual responses from within the different areas generally support the hierarchical model of visual processing [3; 4; 5; 6; 7]. However, both electrophysiological [3; 4; 5] and anatomical [8; 9] studies outline numerous complexities and indeterminacies in the hierarchy and clearly indicate the need for investigation of system dynamics based on single-trial responses.

A prominent basic circuit in the primary visual cortex is the projection from the granular lamina to the supragranular laminae of V1 [10]. The granular layer of V1 (Layer 4) represents the primary thalamorecipient layer of cortex in the visual system, and the supragranular laminae of V1 receive inputs from the granular layer before propagating them to the higher-order visual areas. Both anatomical [2; 10] and electrophysiological data [3; 4; 6] predict a relationship between the amplitudes and latencies of responses simultaneously recorded from these different layers. The mcERP algorithm can address this issue because it estimates the components from the recorded signals and it also estimates their single-trial characteristics. Correlations in the trial-to-trial variability between two components would suggest dynamic coupling between them.

Prior to application of any new model to real data however, it is helpful to characterize the performance of the algorithm with a data set whose elements are known. In this paper, the performance of the mcERP algorithm is shown on numerous synthetic data sets consisting of three physiological components designed to simulate the known features of the granular-supragranular circuit in V1. Each data set contains two hypothetical components that display latency coupling, amplitude coupling, or both latency and amplitude coupling. The accuracy of the algorithm in estimating the three hypothetical components is examined by comparing the estimated parameters of the mcERP model with the original parameters utilized in generating the synthetic data. Also, the implications of these simulations are discussed with respect to

application of the algorithm to real data and with respect to possible future refinements of the model.

METHODS

The mcERP Model and Algorithm

The mcERP model [11; 12; 1] describes recorded neuroelectrophysiological activity as being generated by a set of neural ensembles each responding to stimuli with a stereotypic activation pattern. The stereotypic activity is a temporal pattern of electrical activity that we will describe as a component $s_n(t)$, where s is the component's waveshape across time t , and n is an index for a particular component. Second, amplitude and latency variability in the evoked response occurs from trial to trial [13], and the mcERP model modifies the component waveshape $s_n(t)$ with a trial-specific amplitude scaling factor α_{nr} and a trial-specific latency shift τ_{nr} , where n still denotes a particular component and r denotes a specific trial. And third, since multiple detectors (referred to as electrodes from now on) are often used, a coupling matrix C_{mn} is added to the model to describe the relation between a particular electrode m and a particular component n . This is important since several electrodes may record a particular component differentially. These aspects of the mcERP model can be written mathematically as:

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

where n indexes the N components, C_{mn} denotes the coupling between the m^{th} detector and the n^{th} component, α_{nr} is the amplitude scaling of the n^{th} component during the r^{th} trial, τ_{nr} is the latency shift of the n^{th} component during the r^{th} trial, $s_n(t)$ is the waveshape of the n^{th} component, and $\eta_{mr}(t)$ is the unpredictable signal recorded in the m^{th} detector during the r^{th} trial.

The most probable set of parameters that satisfy the mcERP model expressed in (1) can be calculated from the *maximum a posteriori* (MAP) solution of the posterior probability. The derivation of the equations governing the most probable parameters for (1) begins with Bayes' Theorem, which states that the posterior probability is:

$$p(\text{model} | \text{data}, I) = \frac{p(\text{data} | \text{model}, I) p(\text{model} | I)}{p(\text{data} | I)}, \quad (2)$$

where $p(\bullet)$ is the probability and I is any prior information. The $p(\text{data} | \text{model}, I)$ in (2) is the likelihood of the data given the model, the $p(\text{data} | I)$ is called the evidence or the prior probability of the data, $p(\text{model} | I)$ is called the prior probability of the model, and the left-hand side of (2) is called the posterior probability, which is the probability that a given set of model parameters is consistent with the data and the information. Bayes' Theorem describes how the prior probability of the model

$p(\text{model}|I)$ is modified by new information such as data. Substituting the parameters of the mcERP model into (2), the posterior probability becomes

$$p(\mathbf{C}, \mathbf{s}(t), \mathbf{a}, \boldsymbol{\tau} | \mathbf{x}(t), I) = \frac{p(\mathbf{x}(t) | \mathbf{C}, \mathbf{s}(t), \mathbf{a}, \boldsymbol{\tau}, I) p(\mathbf{C}, \mathbf{s}(t), \mathbf{a}, \boldsymbol{\tau} | I)}{p(\mathbf{x}(t) | I)}. \quad (3)$$

where the bold terms indicate the entire set of a particular parameter. Since the evidence $p(\mathbf{x}(t)|I)$ is constant as the model parameters vary, the posterior probability can be written

$$p(\mathbf{C}, \mathbf{s}(t), \mathbf{a}, \boldsymbol{\tau} | \mathbf{x}(t), I) \propto p(\mathbf{x}(t) | \mathbf{C}, \mathbf{s}(t), \mathbf{a}, \boldsymbol{\tau}, I) p(\mathbf{C}, \mathbf{s}(t), \mathbf{a}, \boldsymbol{\tau} | I) \quad (4)$$

with a proportionality constant equal to the inverse of the evidence. Next, the prior probability of the model can be factored and expressed as the product of the individual probabilities as follows:

$$p(\mathbf{C}, \mathbf{s}(t), \mathbf{a}, \boldsymbol{\tau} | \mathbf{x}(t), I) \propto p(\mathbf{x}(t) | \mathbf{C}, \mathbf{s}(t), \mathbf{a}, \boldsymbol{\tau}, I) p(\mathbf{C} | I) p(\mathbf{s}(t) | I) p(\mathbf{a} | I) p(\boldsymbol{\tau} | I) \quad (5)$$

The $p(\mathbf{a}|I)$ and the $p(\boldsymbol{\tau}|I)$ are assigned uniform distributions and given ranges that are physiologically realizable. A Gaussian likelihood is assigned to (5) based on the principle of maximum entropy [14; 15] resulting in

$$p(\mathbf{C}, \mathbf{s}(t), \mathbf{a}, \boldsymbol{\tau}, \sigma | \mathbf{x}(t), I) \propto (2\pi \sigma^2)^{-\frac{MRT}{2}} \text{Exp}\left[-\frac{1}{2\sigma^2} Q\right] p(\sigma | I) p(\mathbf{C} | I) p(\mathbf{s} | I) \quad (6)$$

where σ is the expected standard deviation between the predictions and the mean, $p(\sigma|I)$ is the prior probability of σ , and Q represents the sum of the square of the residual between the data and model. Q is derived from (1) as

$$Q = \sum_{m=1}^M \sum_{r=1}^R \sum_{t=1}^T \left(x_{mr}(t) - \sum_{n=1}^N C_{mn} \alpha_{nr} s_n(t - \tau_{nr}) \right)^2, \quad (7)$$

where M is the total number of electrodes, R is the number of trial presentations of a stimulus, and T is the number of the data points sampled. A Jeffrey's prior is assigned to σ such that $p(\sigma|I) = \sigma^{-1}$, and the joint posterior probability of (6) is marginalized over all values of σ giving the marginal joint posterior probability

$$p(\mathbf{C}, \mathbf{s}(t), \mathbf{a}, \boldsymbol{\tau} | \mathbf{x}(t), I) \propto Q^{-\frac{MRT}{2}} p(\mathbf{C} | I) p(\mathbf{s} | I). \quad (8)$$

The derivation continues with the assignment of prior probability for the coupling matrix and the component waveshapes. The solution of the electromagnetic forward problem could be used to generate knowledge about the source-detector coupling [16; 17], but for simplicity only uniform prior probabilities were assigned to $p(\mathbf{C}|I)$ and $p(\mathbf{s}|I)$. When is this done and when the natural logarithm is taken, the marginal joint posterior probability becomes

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

where P is $p(\mathbf{C}, \mathbf{s}(t), \boldsymbol{\alpha}, \boldsymbol{\tau} | \mathbf{x}(t), I)$.

Setting the derivative of the posterior probability with respect to each of the model parameters to zero and solving for the parameters yields the *maximum a posteriori* (MAP) solution of (9). These equations are not shown here but are found in Knuth *et al.* [11; 12]. These equations are then incorporated into an iterative algorithm that initially assumes the existence of a single component within the data and refines the parameters $s_1(t)$, α_{1r} , τ_{1r} , and C_{m1} until there is less than 1% change in the waveshape of the $s_1(t)$ or until 15 iterations are complete. After estimating a single component, the residual signal is calculated as

$$residual_{mr}(t) = x_{mr}(t) - \sum_{n=1}^N C_{mn} \alpha_{nr} s_n(t - \tau_{nr}). \quad (10)$$

The residual signals are averaged across trials yielding $\overline{residual}_m(t)$ for each electrode channel, and the current source density (CSD), computed as the second spatial derivative, profile of these residuals is calculated [18; 19; 20]. The CSD indexes the sources (net intracellular to extracellular current flow) and sinks (net extracellular to intracellular current flow) responsible for generating field potentials, and it is utilized routinely to isolate locally generated neuronal activity [20]. If the CSD profile of the average residual signal illustrates evidence of organized sources and sinks similar to those observed empirically [20], we conclude that the residual signal has structure, and the mcERP algorithm is applied to the field potential data again to estimate an additional component $s_2(t)$ and its related parameters. In this second set of iterations, both $s_1(t)$, $s_2(t)$, and their related parameters are all updated. This process of adding components continues until the user determines that the residual signal contains no structure across trials.

The Synthetic Data

Synthetic data were created to evaluate the performance of the algorithm on data containing two interacting neural sources. The structure of the synthetic data simulated field potential recordings from a linear, 15-channel multielectrode array spanning laminae 1 through 6 of macaque V1 during presentation of a red flash of light. In each simulated case, 50 trials of data for each of the 15 channels of the electrode array were generated. Three separate components were specified (Figure 1A): (i) Component 1 resembled activation of a spiny stellate cell population in granular lamina 4 by a thalamic input; (ii) Component 2 represented activation of pyramidal cells in supragranular lamina 2 or 3 by a granular input such as component 1; and (iii) Component 3 is the electrical activity generated by a far-field source. The temporal patterns of these components were chosen to resemble average signals observed in V1 in response to a red flash of light. A coupling matrix specified the relative positioning of the different sources such that component 1 was located between channels 9 and 10, component 2 was located between channels 4 and 5, and component 3 was seen nearly equally across the channels of the electrode array (Figure 1B). Low-frequency noise with a spectrum of $1/f$ (f denotes frequency) was

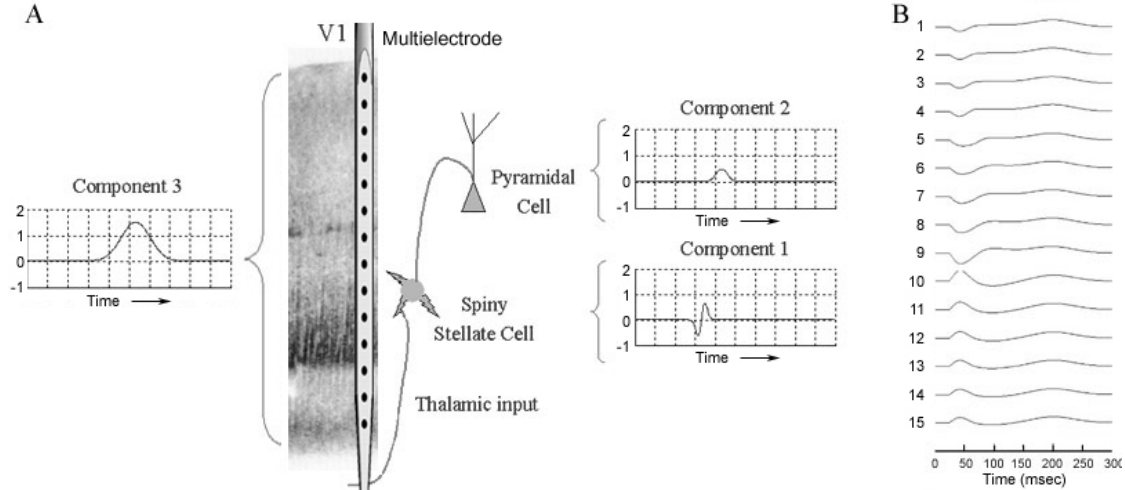


Figure 1. Synthetic data created to evaluate the mCERP algorithm were based on field potential recordings from a linear, multi-electrode array positioned to span laminae 1 through 6 of V1 simultaneously. A histological section of V1 is shown for reference. **A.** Component 1 of the synthetic data represents activation of the lamina 4 spiny stellate cells by a thalamic input. Component 2 represents the subsequent activation of pyramidal cells in the lamina 2 or 3, and component 3 is a far-field source whose component is recorded approximately equally by all electrode channels. **B.** The spatial distribution of the three components is observed in their linear summation at every electrode channel. Component 1 is located between channels 9 and 10, component 2 is located between channels 4 and 5 (difficult to observe here), and component 3 is seen nearly equally by each electrode channel.

added to the data, and the standard deviation of the noise was specified as 0.063 so that the signal-to-noise ratio (SNR) of the smallest component (component 2) was 1.1 as given by the equation $SNR = \sigma_{component}^2 / \sigma_{noise}^2$ where σ denotes the standard deviation across time.

Components 1 and 2 of the synthetic data were coupled so that component 1 drives component 2 via a feedforward connection. We studied four different implementations of this relationship. Method 1 specified a relationship between the latency shifts of components 1 and 2 so that when the first component was delayed, so was the second, and when the first component was early, so was the second. The degree to which the sources were coupled was controlled by a coupling parameter k (not to be confused with the coupling matrix). When k was zero, the sources were uncoupled and the latency of component 2 remained independent of component 1. When k was one, the sources were completely coupled so that component 2 followed component 1 precisely. This can be written in equation form as

$$\tau_{2r} = (1 - k)N(0, \sigma_{jlat}) + k\tau_{1r}, \quad (11)$$

where τ_{2r} is the latency shift of the second component in the r^{th} trial, k is a constant that varies the coupling between the components, $N(0, \sigma_{jlat})$ represents a random number sampled from a normal distribution with mean equal to 0 and standard deviation equal to σ_{jlat} , and τ_{1r} is the latency shift of the first component in the r^{th} trial. Single-trial synthetic data were generated for the 15 electrode channels during 50 trials of the stimulus. The amplitude scaling factors of the three components in each trial were independent of each other and defined from a randomly sampled lognormal

distribution with a sample mean $\mu_{amp} = 1.0$ and sample standard deviation $\sigma_{amp} = 1.0$. The latency shifts of component 1 and component 3 were independent and defined from a normal distribution $N(0, 10.0 \text{ milliseconds})$ with sample mean $\mu_{lat} = 0.0$ milliseconds and sample standard deviation $\sigma_{lat} = 10.0$ milliseconds. Eleven separate cases were explored with $k = \{0.0, 0.1, \dots, 1.0\}$. In each case, σ_{jlat} from (11) was equal to 10.0 milliseconds.

Method 2 specified a coupling between the amplitudes of components 1 and 2. This relation was given by:

$$\alpha_{2r} = (1-k) \times 2^{N(1, \sigma_{jamp})} + k \frac{2}{1 + e^{-b(\alpha_{1r}-1)}}, \quad (12)$$

where α_{1r} and α_{2r} are the amplitude scaling factors of the first and second components respectively in the r^{th} trial, k is a constant that varies the coupling, and $N(1, \sigma_{jamp})$ represents a random number from a normal distribution with sample mean equal to 0 and sample standard deviation equal to σ_{jamp} . The first term in the sum on the right-hand side of (12) represents the independent lognormal amplitude jitter of component 2, and the second term represents a sigmoid function whose rise is varied by b (Figure 2). As in Method 1, single-trial data for 15 electrode channels were generated across 50 trials. The latency shifts of the three components were independent and chosen randomly from $N(0, 10.0 \text{ milliseconds})$. The amplitude scaling factors of components 1 and 3 were also independent and chosen from a lognormal distribution with a sample mean $\mu_{amp} = 1.0$ and sample standard deviation $\sigma_{amp} = 1.0$. Forty different simulations were performed using this method with $\sigma_{jamp} = 0.75$, $b = \{2^0, 2^1, 2^2, 2^3\}$, and $k = \{0.1, 0.2, \dots, 1.0\}$ for each value of b . Since a negative value for α is not physiologically

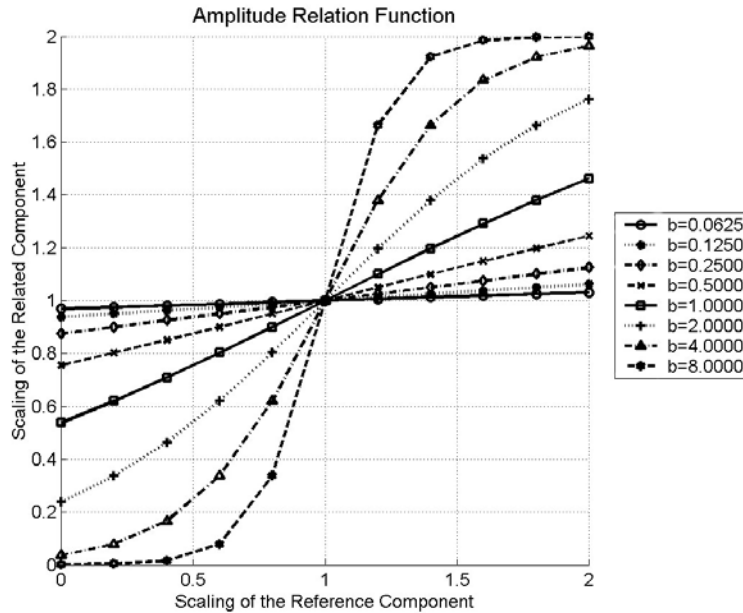


Figure 2. Amplitude relation function used in the simulated cases studying amplitude relation alone and those studying simultaneous amplitude and latency relation. As the value of b is increased, the relation varies from linear to sigmoidal.

realizable, each α_{2r} value was adjusted by $(1 - \langle \alpha_2 \rangle)$, where $\langle \alpha_2 \rangle$ represents the mean of all amplitude scaling factors across R trials.

Method 3 combined elements of the two previous methods so that the components were coupled both by latency shifts and amplitude scaling factors. The latency and amplitude relations were specified by (11) and (12) respectively, and the single-trial data were again generated for 15 electrode channels during 50 trial presentations of the stimulus with the single-trial amplitudes and latencies of components 1 and 3 determined as before. The latency and amplitude relationships between components 1 and 2 were calculated with $\sigma_{jlat} = 10.0$ milliseconds, $\sigma_{jamp} = 0.75$, $b = \{2^{-4}, 2^{-3}, 2^{-2}, 2^{-1}, 2^0, 2^1, 2^2, 2^3\}$, and $k = \{0.0, 0.1, \dots, 1.0\}$ for each value of b resulting in 77 simulated cases.

Finally, Method 4 was identical to Method 3 except that the stereotypic waveshape of component 2 was altered to be identical to that of component 1. The latency and amplitude relations were specified with $\sigma_{jlat} = 10.0$ milliseconds, $\sigma_{jamp} = 0.75$, $b = \{2^{-3}, 2^{-1}, 2^0, 2^1, 2^2, 2^3\}$, and $k = \{0.0, 0.1, \dots, 1.0\}$ for each value of b . Thus sixty-six different combinations were simulated.

Evaluation of Algorithm Performance

The mcERP algorithm's performance was evaluated with several different measures of accuracy. First, the Amari error [21] was calculated to quantify the algorithm's ability to separate each component comprising the synthetic data. The basis of this error is derived from the linear mixing model given by

$$x = Cs, \quad (13)$$

where x is a matrix of the signals recorded in each detector across time, C is the mixing (or coupling) matrix describing the component-detector relationship, and s is a matrix of the component waveshapes across time. The linear mixing model of (13) possesses two indeterminacies because the coupling coefficients in matrix C and the component waveshapes in matrix s can be rescaled by a diagonal scaling matrix Σ such that

$$x = C \Sigma \Sigma^{-1} s, \quad (14)$$

and the order of the sources can arbitrarily be changed by a permutation matrix Π so that

$$x = C \Sigma \Pi \Pi^{-1} \Sigma^{-1} s. \quad (15)$$

Thus, the linear mixing model can be rewritten as

$$x = (C \Sigma \Pi) (\Pi^{-1} \Sigma^{-1} s). \quad (16)$$

Expression (16) asserts that estimates of the coupling matrix and the component waveshapes will be as accurate as a scaled permutation of the original and may be written as

$$\hat{s} = M^{-1} s. \quad (17)$$

or

$$\hat{C} = CM. \quad (18)$$

where M is a matrix describing the relation between the original matrix and the estimated version. A perfect estimate of the coupling matrix or the component waveshapes would result in M being a scaled permutation matrix given by $M = \Sigma II$. Thus, studying M yields insight into the accuracy of the estimated quantity. Solving for M^{-1} in (17) gives

$$M^{-1} = (\hat{s}s^T)^{-1} \quad (19)$$

and in (18) yields

$$M^{-1} = \left((C^T C)^{-1} (C^T \hat{C}) \right)^{-1} = (C^T \hat{C})^{-1} (C^T C), \quad (20)$$

where T denotes the transpose of the marked matrix. The deviation of M^{-1} from a scaled permutation matrix is found by computing the Amari error

$$E_{Amari} = \frac{\sum_{i=1}^n \left(\frac{\sum_{j=1}^n |A_{ij}|}{\max_k |A_{ik}|} - 1 \right) + \sum_{j=1}^n \left(\frac{\sum_{i=1}^n |A_{ij}|}{\max_k |A_{kj}|} - 1 \right)}{2(n^2 - n)}, \quad (21)$$

where A was set equal to M^{-1} , i indexes the rows of matrix A , j indexes the columns of matrix A , n is the total number of components, k is the index of the position of the maximal value, and the denominator normalizes the error. Either (19) or (20) can be used to calculate the Amari error; (19) was used to calculate the Amari error for relationship methods 1, 2, and 3, and (20) was used to evaluate the Amari error in relationship method 4 because the quantity $(ss^T)^{-1}$ is singular when components 1 and 2 have the same waveshape.

Comparing the estimated component waveshapes to the actual component waveshapes was the second tool used to evaluate the performance of the algorithm. The root-mean-square (RMS) error was calculated as follows:

$$E_{wave}^j = \frac{\sqrt{\sum_{t=1}^T (s_j(t) - \hat{s}'_j(t))^2}}{\sqrt{\sum_{t=1}^T (s_j(t))^2}}, \quad (22)$$

where j indexes a particular component, T is the total number of time samples, and \hat{s}' denotes the estimated sources \hat{s} after they have been scaled and permuted so as to be comparable to the original sources.

Accuracy of the latency shifts and amplitude scaling factors was examined by calculating the average absolute difference (average ABS error) between the original values and the estimated ones. The following two equations define these errors for each component j across all R trials:

$$E_{amp}^j = \frac{1}{R} \sum_{r=1}^R |\alpha_{jr} - \hat{\alpha}_{jr}|; \quad (23)$$

$$E_{lat}^j = \frac{1}{R} \sum_{r=1}^R |\tau_{jr} - \hat{\tau}_{jr}|. \quad (24)$$

ABS errors below the standard deviation of the variability ($\sigma_{jamp}=1.0$ and $\sigma_{jlat}=10.0$ milliseconds) signify an improvement in estimation over that of standard averaging.

RESULTS

Performance was first evaluated for synthetic data generated using coupling method 1. Each measure of performance showed that a strong latency coupling alone does not hamper mcERP’s ability to estimate the parameters. For each value of k , the Amari error (Figure 3A) remains below 0.06, which indicates less than about 1% error in separation. Although not shown, the component RMS error is below 10% and does not illustrate a trend with increasing coupling. Finally the average absolute errors of the amplitude scaling factors and the latency shifts are well below the standard deviations of those quantities in the original data.

Method 2 related the amplitude of component 2 to the amplitude of component 1 via a sigmoid function (12) with a parameter b (Figure 2), which controls the slope of the transition (larger b gives a larger slope). Both the transition parameter b and the coupling constant k were varied. Figure 3B illustrates the Amari error for method 2. Again the Amari errors were relatively constant as a function of b and k , and were less than 0.06. The other parameters were also well estimated.

Method 3 combined the aspects of methods 1 and 2 such that simultaneous amplitude and latency relations were considered. Again, the mcERP algorithm estimated the components and their single-trial parameters accurately. Figure 4A shows that the Amari error does not vary across k or b .

The final test of the mcERP algorithm studied synthetic data whose component 1 and component 2 had identical waveshapes and were related both in amplitude and latency. Table 1 lists the Amari errors for all of the different cases considered in this block of simulations; there are only five cases where the Amari error indicates less than 1 percent error in separation. Figure 4B illustrates a surface plot of these values, and shows that as strength of the interaction k increases, the Amari error also increases. In contrast, as the amplitude interaction varies from roughly linear (small b)

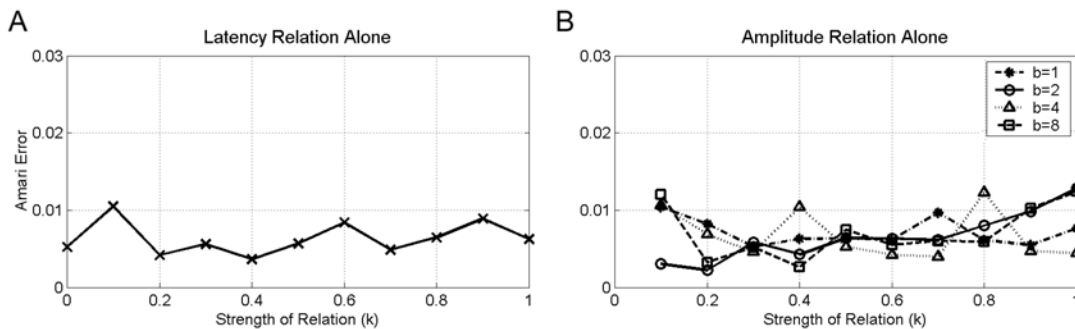


Figure 3. The resulting Amari error when the mcERP algorithm is applied to synthetic data with a latency relation alone (A) and with an amplitude relation alone (B).

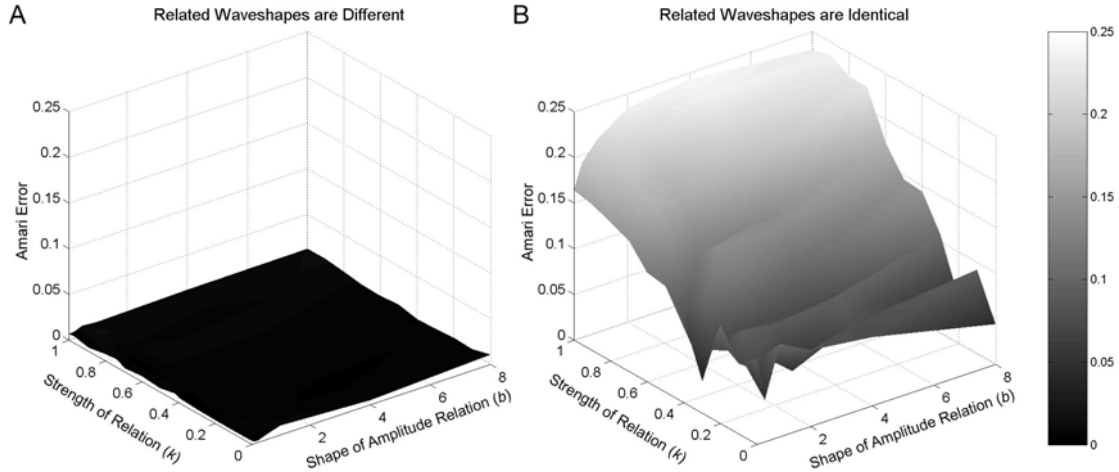


Figure 4. mcERP algorithm performance on synthetic data containing both an amplitude and a latency relation between components 1 and 2. Panel **A** shows the Amari error when the related components' waveshapes differed, and **B** shows the Amari error when the components' waveshapes were identical.

to sigmoid-shaped (large b), the Amari error does not follow an appreciable pattern. Examining the component waveshape, amplitude scaling factor, and latency shift errors (all not shown), several different relationships were observed. First, component 1 waveshape error increased with increasing k while remaining relatively constant with varying b . Component 1's estimated amplitude scaling factor error was largest when both k and b were small, and its estimated latency shift error decreased with increasing k while remaining constant with varying b . Second, component 2's RMS waveshape error, average ABS amplitude scaling factor error, and average ABS latency shift error all increased with increasing k and all increased slightly with increasing b . Finally, the accuracies of component 3 estimates (not shown) exhibit no relation to both k and b .

Samples of the results for two different cases are shown in Figure 5 and Figure 6. Figure 5A-C illustrates an example from method 3 where the waveshapes of components are different, and where $b = 8.0$ and $k = 1.0$. The estimated component waveshapes, their spatial distributions, amplitude scaling factors, and latency shifts are

Table 1. Amari error when the mcERP algorithm is applied to synthetic data where components 1 and 2 have the same waveshape and an amplitude and latency relation.

k value	b value					
	0.2500	0.50	1.00	2.00	4.00	8.00
0.0	0.0934	0.0461	0.1015	0.0835	0.0749	0.0438
0.1	0.0773	0.0714	0.0765	0.0491	0.0790	0.0923
0.2	0.1081	0.0771	0.0942	0.0888	0.0966	0.0482
0.3	0.0350	0.0662	0.0738	0.0744	0.0746	0.1077
0.4	0.0931	0.0605	0.1075	0.1205	0.1037	0.1431
0.5	0.1160	0.0949	0.0930	0.1407	0.1379	0.1452
0.6	0.1200	0.1286	0.1399	0.1226	0.1360	0.1742
0.7	0.1431	0.1574	0.1737	0.1662	0.1840	0.2227
0.8	0.1521	0.1664	0.1848	0.1805	0.2104	0.2241
0.9	0.1593	0.1828	0.1968	0.2098	0.2360	0.2350
1.0	0.1641	0.1908	0.2112	0.2342	0.2428	0.2298

shown with respect to their original counterparts. Figure 6A-C shows a case from relation method 4 where the waveshapes of components 1 and 2 were identical, and where $b = 8.0$ and $k = 1.0$. In this case, the component waveshapes and their spatial distributions are not well estimated because the components are mixed. In addition, method 4 results in greater errors in the estimates of the single-trial features as observed in Figure 6B and C in comparison to those illustrated in Figure 5B and C.

DISCUSSION

The mcERP model describes the single-trial event-related potential as the linear summation of multiple components each described by a stereotypic waveshape and possessing a trial-specific amplitude and latency. Application of Bayes' theorem to the model allows development of an algorithm that gives the most probable set of parameters to satisfy the model. The advantage of utilizing Bayes' theorem is that algorithm failure results only from model inadequacies, inappropriate or insufficient prior information about the situation, or insufficient data. In many cases, model inadequacies may be identified by testing the algorithm on synthetic data prior to application to real data.

Dynamical coupling between real components is expected, and accurately indexing this coupling will yield insights into sensory processing within and between cortical areas. We must first evaluate the performance of the algorithm on data with interacting components by testing it on numerous synthetic data sets that reconstruct four hypothetical relations between granular and supragranular components evoked by an external stimulus. Amplitude-amplitude and latency-latency interactions between two components were found not to affect the ability of the algorithm to identify the component waveshapes, their spatial locations, or their single-trial amplitude and latency characteristics. These latency-latency results were expected as previous studies of the mcERP algorithm found that accurate identification of three components with independent amplitude variability but no latency variability was possible [12]. This is similar to method 1 presented here because a lack of latency variability in the components can be described by (11) when $\tau_{lr} = 0$.

The second set of simulations explored an amplitude interaction between two components. A sigmoid relation was utilized because it provides a threshold and saturation. The threshold of the sigmoid relation prevents occurrence of the driven component when the driving component's activation level is below a given threshold. The saturation of the sigmoid curve prevents a further increase in the driven component's amplitude when all elements of the neuronal ensemble generating this component are "recruited." In this set of simulations, the algorithm accurately estimated all model parameters.

Simultaneous amplitude and latency interactions between the granular and the supragranular components were examined in the third group of simulations. Again, the mcERP algorithm accurately estimated all parameters. We hypothesized that although the single-trial characteristics of components 1 and 2 were related, mcERP succeeded in the identification because the waveshapes of the components were

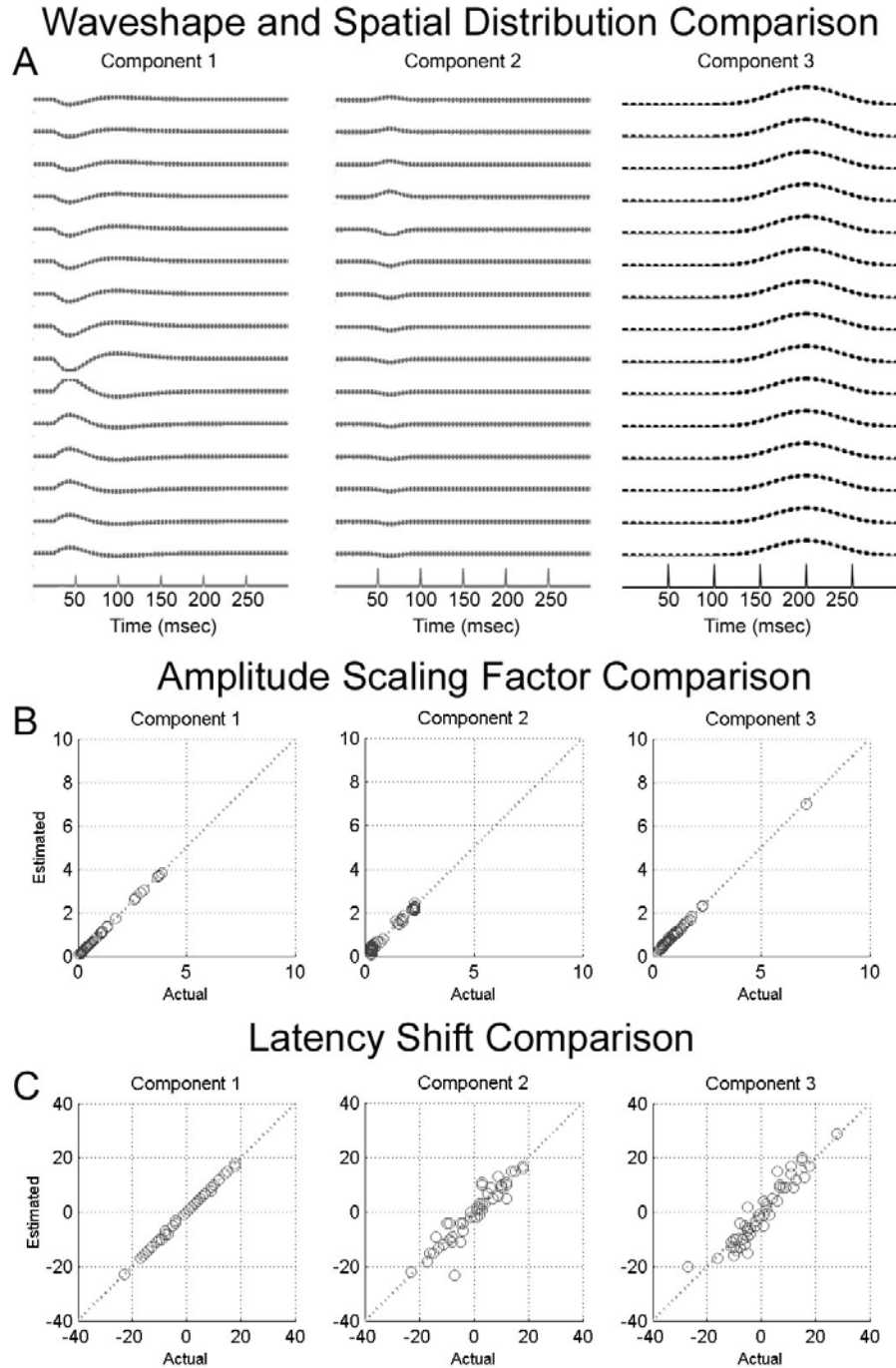


Figure 5. mcERP algorithm estimates match the original synthetic data well. In this case, the waveshapes of components 1 and 2 are different, $k = 1.0$, and $b = 8.0$. **A.** The component waveshapes and their spatial distributions are plotted across electrode channels. The dotted lines in each plot indicate the estimated component, and the solid lines indicate the original waveshapes. The solid lines are difficult to observe because the estimated components match the original ones well. **B.** Comparison of the actual and estimated amplitude scaling factors. **C.** Comparison of the actual and estimated latency timeshifts. The dotted diagonal line denotes a perfect match between the actual and the estimated values in both B and C.

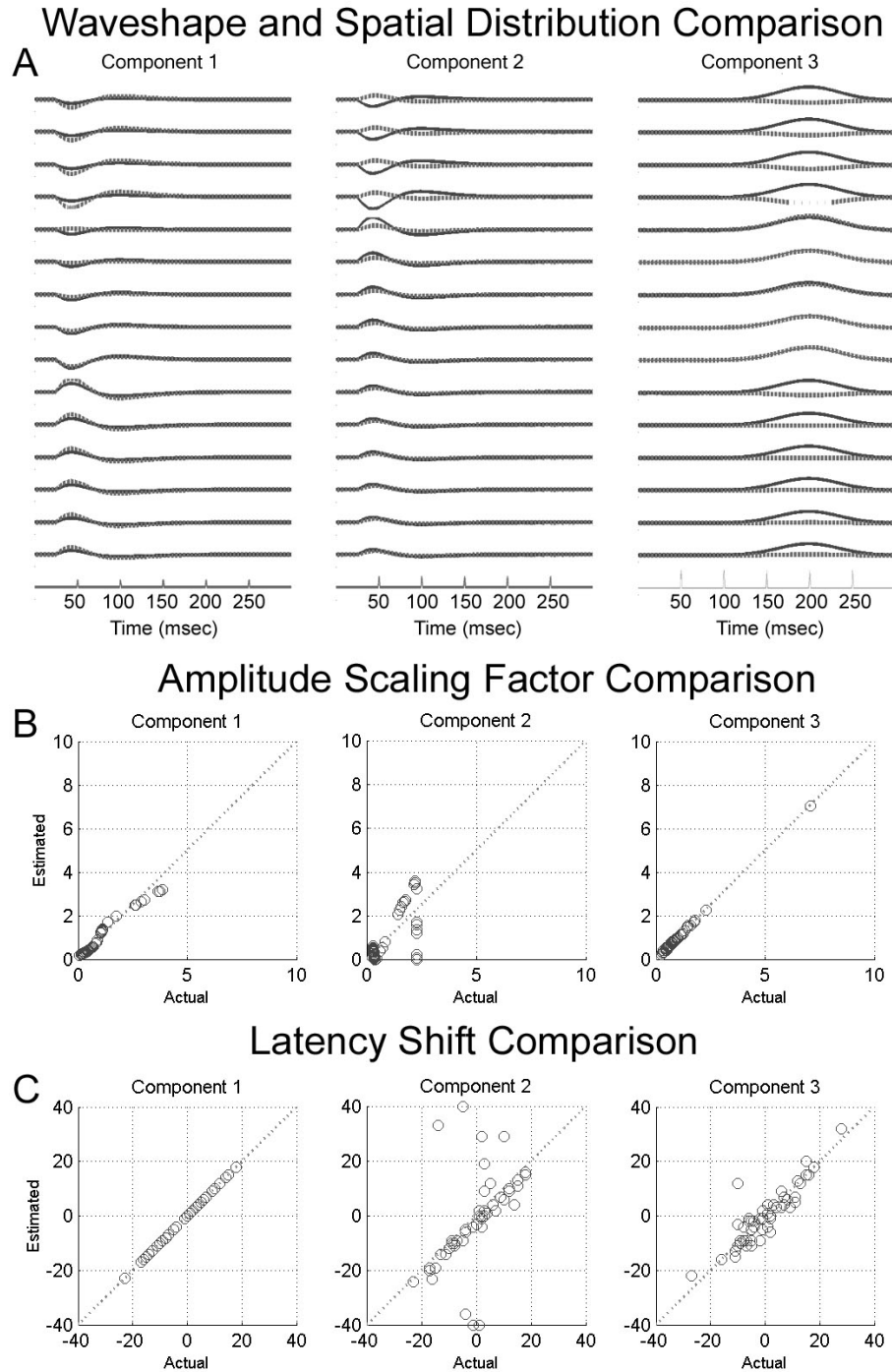


Figure 6. mcERP algorithm estimates do not match the original synthetic data well. In this particular case, the waveshapes of components 1 and 2 are equivalent, $k = 1.0$, and $b = 8.0$. **A.** The component waveshapes and their spatial distributions are plotted across electrode channels. The dashed lines indicate the estimated component, and the solid lines indicate the original waveshapes. **B.** Comparison of the actual and estimated amplitude scaling factors. **C.** Comparison of the actual and estimated latency timeshifts. The dotted, diagonal line denotes a perfect match between the actual and the estimated values in both B and C.

different. To test this hypothesis, we performed the fourth set of simulations where the waveshapes of the coupled components were identical. In these simulations, the quality of separation, as measured by the Amari error, degrades as the strength of their interaction increases, but it remains relatively unaffected by changes in the shape of the amplitude relation. The amplitude scaling factor error and the latency shift error show an inverse relationship between the two related components. Since there is considerable mixing between these two components, this observation may indicate that the first of the two related components estimated by the algorithm is better approximated than the second. Indeed, the waveshapes for the highly coupled case in Figure 6 suggest that this is the case. Together these findings indicate that a difference in the temporal activation pattern of the components is crucial in separating strongly interacting components. When the component waveshapes are the same, independent variation of the amplitude and latency parameters of the components can assist the algorithm in separating the components, but this is not typically sufficient to accurately estimate all of the parameters.

Although several scenarios were considered, other interactions need to be investigated. For example, amplitude-latency couplings may exist so that a larger than average amplitude of a component in a given trial may result in the earlier activation of a second, dynamically coupled component. This scenario has not been explored fully. Also, the effect of the spatial distribution on the estimation of two dynamically coupled sources will require further study. Lastly, the degree to which differences in the waveshapes affect the quality of separation remains to be seen.

By applying the mcERP algorithm to real data we expect to be able to characterize the single-trial properties of these evoked responses, which will provide valuable insights into the cortical processing of sensory stimuli. The results of the simulations presented here suggest that the mcERP algorithm will be able to distinguish between spatially distinct neural ensembles generating differing component waveshapes despite the fact that they may be interacting with one another. However, as the waveshapes of the two components become more similar, we expect the quality of separation to degrade. This scenario may be addressed by incorporating more prior information into the mcERP algorithm by adopting a sparse prior on the coupling matrix to enforce spatial localization of the sources. Such prior information would also offset the effects of real variations in the recording apparatus such as impedance mismatches among electrodes. Lastly, we currently employ a crude and rather subjective stopping criterion, which could be substantially improved by deriving a criterion that is based on the principles of model selection. This would avoid both over- and under-fitting of the data. Regardless, the mcERP algorithm is performing as expected and the results of these simulations give us confidence that the algorithm will produce viable results when applied to real data.

ACKNOWLEDGMENTS

Supported in part by T32M07288, MH60358, NARSAD Young Investigator Award, IBN0090717, and MH64204. We thank Dr. Jonathan Victor and his laboratory for numerous helpful discussions, and Tammy McGinnis and Noelle

O’Connell for technical support. We also thank our colleagues Kaiming Fu, Taylor Johnston, Dr. Donald Faber, Dr. Alan Finkelstein, and Dr. Mitchell Steinschneider for helpful comments on an earlier version of the model.

REFERENCES

1. Truccolo, W.A., Knuth, K.H., Ding, M., and Bressler, S.L., “Bayesian estimation of amplitude, latency and waveform of single trial cortical evoked components,” in *Maximum Entropy and Bayesian Methods, Baltimore 2001*, edited by R.L. Fry and M. Bierbaum, American Institute of Physics, Melville, New York, 2002, pp. 64-72.
2. Felleman, D.J. and Van Essen, D.C., *Cerebral Cortex*, **1(1)**, 1-47 (1991).
3. Schroeder, C.E., Mehta, A.D., and Givre, S.J., *Cerebral Cortex*, **8(7)**, 575-92 (1998).
4. Givre, S.J., Schroeder, C.E., and Arezzo, J.C., *Vision Research*, **34(4)**, 415-28 (1994).
5. Nowak, L.G. and Bullier, J., “The Timing of Information Transfer in the Visual System,” in *Cerebral Cortex*, vol 12, edited by Rockland *et al.*, Plenum Press, New York, 1997, pp. 205-241.
6. Maunsell, J.H. and Gibson, J.R., *Journal of Neurophysiology*, **68(4)**, 1332-44 (1992).
7. Schmolesky, M.T., Wang, Y., Hanes, D.P., Thompson, K.G., Leutgeb, S., Schall, J.D., and Leventhal, A.G., *Journal of Neurophysiology*, **15(6)**, 1107-18 (1998).
8. Hilgetag, C.C., O’Neill, M.A., and Young, M.P., *Science*, **271(5250)**, 776-7 (1996).
9. Kotter, R. and Sommer, F.T., *Philos Trans R Soc Lond B Biol Sci*, **355(1393)**, 127-34 (2000).
10. Lund, J.S., *Annual Revue of Neuroscience*, **11**, 253-288 (1988).
11. Knuth, K.H., Truccolo, W.A., Bressler, S.L., and Ding, M., “Separation of Multiple Evoked Responses Using Differential Amplitude and Latency Variability,” in *Proceedings of the Third International Workshop on Independent Component Analysis and Blind Signal Separation: ICA 2001*, San Diego, California, 2001.
12. Knuth, K.H., Shah, A.S., Truccolo, W.A., Ding, M., Bressler, S.L., and Schroeder, C.E., “Multiple Component Event-Related Potential (mcERP) Estimation using Differential Amplitude and Latency Variability,” *in preparation*.
13. Truccolo, W.A., Ding, M., Knuth, K.H., Nakamura, R., and Bressler, S.L., *Clinical Neurophysiology*, **113(2)**, 206-26 (2002).
14. Sivia, D.S., *Data Analysis. A Bayesian Tutorial*, Clarendon Press, Oxford, 1996, pp. 113-120.
15. Jaynes, E.T. *Probability Theory - The Logic of Science*, <http://bayes.wustl.edu>, unpublished.
16. Knuth, K.H., “Bayesian source separation and localization”, in *Bayesian Inference for Inverse Problems*, edited by A. Mohammad-Djafari, Proceedings of SPIE Volume 3459, San Diego, California, 1998, pp. 147-58.
17. Knuth, K.H. and Vaughan, H.G. Jr., “Convergent Bayesian formulations of blind source separation and electromagnetic source estimation”, in *Maximum Entropy and Bayesian Methods, Garching, Germany 1998*, edited by W. von der Linden, V. Dose, R. Fischer, and R. Preuss, Kluwer Academic Publishers, Dordrecht, 1999, pp. 217-26.
18. Freeman, J.A. and Nicholson, C., *Journal of Neurophysiology*, **38**, 369-382 (1975).
19. Mitzdorf, U., *Physiol Reviews*, **65**, 37-100 (1985).
20. Schroeder, C.E., *Int J Neurosci*, **80**, 65-78 (1995).
21. Amari, S., Cichocki, A., and Yang, H.H., “A New Learning Algorithm for Blind Signal Separation,” in *Advances in Neural Information Processing Systems 8*, edited by D. Touretzky, M. Mozer, and M. Hasselmo, MIT Press, Cambridge, Massachusetts, 1996, pp. 752-763.