State-space modeling for single-trial evoked potential estimation

Stefanos Georgiadis, Perttu Ranta-aho, Mika Tarvainen and Pasi Karjalainen
Department of Physics, University of Kuopio, Kuopio
Finland

1. Introduction

The exploration of brain responses following environmental inputs or in the context of dynamic cognitive changes is crucial for better understanding the central nervous system (CNS). However, the limited signal-to-noise ratio of non-invasive brain signals, such as evoked potentials (EPs), makes the detection of single-trial events a difficult estimation task. In this chapter, focus is given on the state-space approach for modeling brain responses following stimulation of the CNS.

Many problems of fundamental and practical importance in science and engineering require the estimation of the state of a system that changes over time using a series of noisy observations. The state-space approach provides a convenient way for performing time series modeling and multivariate non stationary analysis. Focus is given on the determination of optimal estimates for the state vector of the system. The state vectors provide a description for the dynamics of the system under investigation. For example, in tracking problems the states could be related to the kinematic characteristics of the moving object. In EP analysis, they could be related to trend-like changes of some component of the potentials caused by sequential stimuli presentation. The observation vectors represent noisy measurements that provide information about the state vectors.

In order to analyze a dynamical system, at least two models are required. The first model describes the time evolution of the states, and the second connects observations and states. In the Bayesian state-space formulation those are given in a probabilistic form. For example, the state is assumed to be influenced by unknown disturbances modeled as random noise. This provides a general framework for dynamic state estimation problems. Often, an estimate of the state of the system is required every time that a new measurement is available. A recursive filtering approach is then needed for estimation. Such a filter consists of essentially two stages: prediction and update. In the prediction stage, the state evolution model is used to predict the state forward from one measurement time to the next. The update stage uses the latest measurement to modify the prediction. This is achieved by using the Bayes theorem, which can be seen as a mechanism for updating knowledge about the current state in the light of extra information provided from new observations. When all the measurements are available, that is, in the case of batch processing, then a smoothing strategy is preferable. The smoothing problem can also be treated within the same framework. For example, a forward-

backward approach can be adopted, which gives the smoother estimates as corrections of the filter estimates with the use of an additional backward recursion.

A mathematical way to describe trial-to-trial variations in evoked potentials (EPs) is given by state-space modeling. Linear estimators optimal in the mean square sense can be obtained with the use of Kalman filter and smoother algorithms. Of importance is the parametrization of the problem and the selection of an observation model for estimation. Aim in this chapter is the presentation of a general methodology for dynamical estimation of EPs based on Bayesian estimation theory.

The rest of the chapter is organized as follows: In Section 2, a brief overview of single-trial analysis of EPs is given focusing on dynamical estimation methods. In Section 3, state-space mathematical modeling is presented in a generalized probabilistic framework. In Sections 4 and 5, the linear state-space model for dynamical EP estimation is considered, and Kalman filter and smoother algorithms are presented. In Section 6, a generic way for designing an observation model for dynamical EP estimation is presented. The observation model is constructed based on the impulse response of an FIR filter and can be used for different kind of EPs. This form enables the selection of observation model based on shape characteristics of the EPs, for instance, smoothness, and can be used in parallel with Kalman filtering and smoothing. In Section 7, two illustrative examples based on real EP measurements are presented. It is also demonstrated that for batch processing the use of the smoother algorithm is preferable. Fixed-interval smoothing improves the tracking performance and reduces greater the noise. Finally, Section 8 contains some conclusions and future research directions related to the presented methodology.

2. Single-trial estimation of evoked potentials

Electroencephalogram (EEG) provides information about neuronal dynamics on a millisecond scale. EEG's ability to characterize certain cognitive states and to reveal pathological conditions is well documented (Niedermeyer & da Silva, 1999). EEG is usually recorded with Ag/AgCl electrodes. In order to reduce the contact impedance between the electrode-skin interface, the skin under the electrode is abraded and a conducting electrode past is used. The electrode placement commonly conforms the international 10-20 system shown in Figure 1, or some extensions of it for additional EEG channels. For the names of the EEG channels the following letters are usually used: A = ear lobe, C = central, Pg = nasopharyngeal, P = parietal, F = frontal, Fp = frontal polar, and O = occipital.

Evoked potentials obtained by scalp EEG provide means for studying brain function (Niedermeyer & da Silva, 1999). The measured potentials are often considered as voltage changes resulted by multiple brain generators active in association with the eliciting event, combined with noise, which is background brain activity not related to the event. Additionally, there are contributions from non-neural sources, such as muscle noise and ocular artifacts. In relation to the ongoing EEG, EPs exhibit very small amplitudes, and thus, it is difficult to be detected straight from the EEG recording. Therefore, traditional research and analysis requires an improvement of the signal-to-noise ratio by repeating stimulation, considering unchanged experimental conditions, and finally averaging time locked EEG epochs. It is well known that this signal enhancement leads to loss of information related to trial-to-trial variability (Fell, 2007; Holm et al., 2006).

The term event-related potentials (ERPs) is also used for potentials that are elicited by cognitive activities, thus differentiate them from purely sensory potentials (Niedermeyer & da Silva,
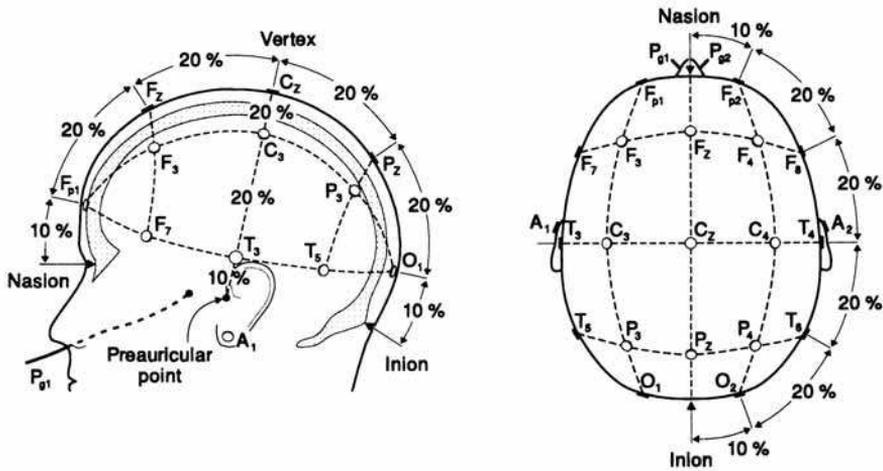
Fig. 1. The international 10-20 electrode system, redrawn from (Malmivuo & Plonsey, 1995).

1999). A generally accepted EP terminology denotes the polarity of a detected peak with the letter "N" for negative and "P" for positive, followed by a number indicating the typical latency. For example, the P300 wave is an ERP seen as a positive deflection in voltage at a latency of roughly 300 ms in the EEG. In practice, the P300 waveform can be evoked using a stimulus delivered by one of the sensory modalities. One typical procedure is the oddball paradigm, whereby a deviant (target) stimulus is presented amongst more frequent standard background stimuli. Elicitation of P300 type of responses usually requires a cognitive action to the target stimuli by the test subject. An example of traditional EP analysis, that is averaging epochs sampled relative to the two types of stimuli, here involving auditory stimulation, is presented in Figure 2. In Figure 2 (a) it is shown the extraction of time-locked EEG epochs from continuous measurements from an EEG channel. In this plot, markers (+) indicate stimuli presentation time. In Figure 2 (b), the average responses for standard and deviant stimuli are presented, and zero at the x-axis indicates stimuli presentation time. Notice, that often the potentials are plotted in reverse polarity.

Evoked potentials are assumed to be generated either separately of ongoing brain activity, or through stimulus-induced reorganization of ongoing activity. For example, it might be possible that during the performance of an auditory oddball discrimination task, the brain activity is being restructured as attention is focused on the target stimulus (Intriligator & Polich, 1994). Phase synchronization of ongoing brain activity is one possible mechanism for the generation of EPs. That is, following the onset of a sensory stimulus the phase distribution of ongoing activity changes from uniform to one which is centered around a specific phase (Makeig et al., 2004). Moreover, several studies have concluded that averaged EPs are not separate from ongoing cortical processes, but rather, are generated by phase synchronization and partial phase-resetting of ongoing activity (Jansen et al., 2003; Makeig et al., 2002). Though, phase coherence over trials observed with common signal decomposition methods (e.g. wavelets) can result both from a phase-coherent state of ongoing rhythms and from the presence of

a phase-coherent EP which is additive to ongoing EEG (Makeig et al., 2004; Mäkinen et al., 2005). Furthermore, stochastic changes in amplitude and latency of different components of the EPs are able to explain the inter trial variability of the measurements (Knuth et al., 2006; Mäkinen et al., 2005; Truccolo et al., 2002). Perhaps both type of variability may be present in EP signals (Fell, 2007).

Several methods have been proposed for EP estimation and denoising, e.g. (Cerutti et al., 1987; Delorme & Makeig, 2004; Karjalainen et al., 1999; Li et al., 2009; Quiroga & Garcia, 2003; Ranta-aho et al., 2003). The performance and applicability of every single-trial estimation method depends on the prior information used and the statistical properties of the EP signals. In general, the exploration of single-trial variability in event related experiments is critical for the study of the central nervous system (Debener et al., 2006; Fell, 2007; Makeig et al., 2002). For example, single-trial EPs could be used to study perceptual changes or to reveal complicated cognitive processes, such as memory formation. Here, we focus on the case that some parameters of the EPs change dynamically from stimulus-to-stimulus. This situation could be a trend-like change of the amplitude or latency of some EP component.

The most obvious way to handle time variations between single-trial measurements is sub-averaging of the measurements in groups. Sub-averaging could give optimal estimators if the EPs are assumed to be invariant within the sub-averaged groups. A better approach is to use moving window or exponentially weighted average filters, see for example (Delorme & Makeig, 2004; Doncarli et al., 1992; Thakor et al., 1991). A few adaptive filtering methods have also been proposed for EP estimation, especially for brain stem potential tracking, e.g. (Qiu et al., 2006). The statistical properties of some moving average filters and different recursive estimation methods for EP estimation have been discussed in (Georgiadis et al., 2005b). Some smoothing methods have also been proposed for modeling trial-to-trial variability in EPs (Turetsky et al., 1989). Kalman smoother algorithm for single-trial EP estimation was introduced in (Georgiadis et al., 2005a), see also (Georgiadis, 2007; Georgiadis et al., 2007; 2008).

State-space modeling for single-trial dynamical estimation considers the EP as a vector valued random process with stochastic fluctuations from stimulus-to-stimulus (Georgiadis et al., 2005b). Then past and future realizations contain information of relevance to be used in the estimation procedure. Estimates for the states, that are optimal in the mean square sense, are given by Kalman filter and smoother algorithms. Of importance is the parametrization of the problem and the selection of an observation model for the measurements. For example, in (Georgiadis et al., 2005b; Qiu et al., 2006) generic observation models were used based on time-shifted Gaussian smooth functions. Furthermore, data based observation models can also be used (Georgiadis, 2007).

3. Bayesian formulation of the problem

In this chapter, sequential observations are considered to be available at discrete time instances t. The observation vector z_t is assumed to be related to some unobserved parameter vector (state vector) through some model of the form
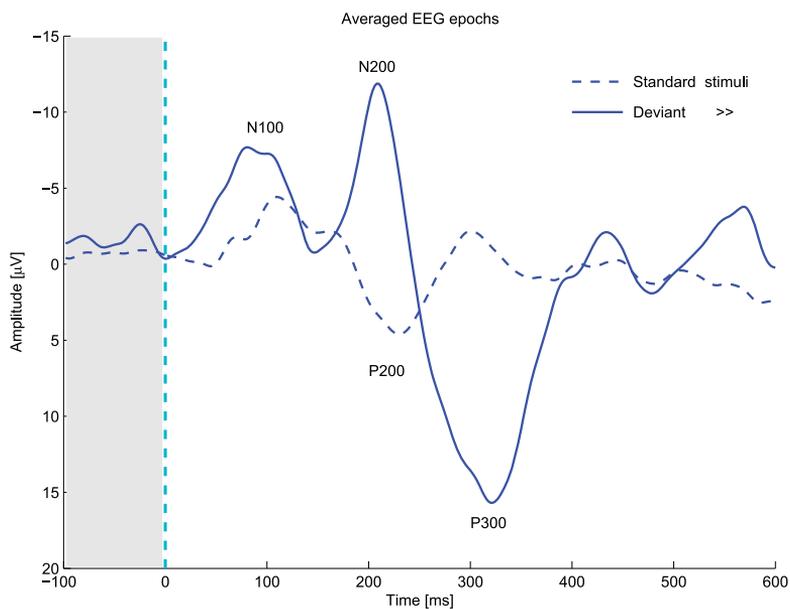
$$z_t = h_t(\theta_t, v_t), \tag{1}$$

for every $t = 1, 2, \ldots$. The simplest non stationary process that can serve as a model for the time evolution of the states is the first order Markov process. This can be expressed with the following state equation:

$$\theta_t = f_t(\theta_{t-1}, \omega_t). \tag{2}$$
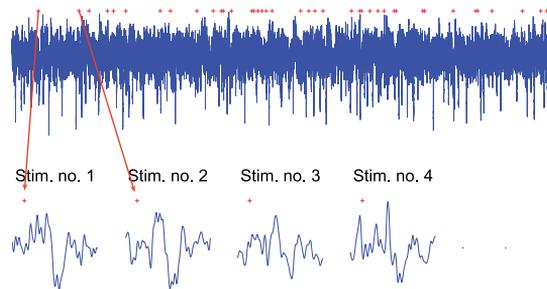
(a) Extracting EEG epochs.

(b) Comparing the average responses.

Fig. 2. Traditional EP analysis for a stimuli discrimination task.

The last two equations form a state-space model for estimation. Other common assumptions made for the model are summarized bellow:

- f_t, h_t are well defined vector valued functions for all t.
- $\{\omega_t\}$ is a sequence of independent random vectors with different distributions, and represents the state noise process.
- $\{v_t\}$ is a white noise vector process, that represents the observation noise.

- The random vectors ω_t, v_t are mutually independent for every t.

- The distributions of ω_t, v_t are known or preselected.

- There is an initial state θ_0 with known distribution.

The previous estimation problem can also be described in a different way. The stochastic process $\{\theta_t\}, \{z_t\}$ are said to form a (first order) evolution-observation pair, if for some random starting point θ_0 and some evolution up to t the following properties hold (Kaipio & Somersalo, 2005):

- The process $\{\theta_t\}$ is a Markov process, that is,

$$p(\theta_t|\theta_{t-1}, \theta_{t-2}, \ldots, \theta_0) = p(\theta_t|\theta_{t-1}). \tag{3}$$

- The process $\{z_t\}$ has the memory-less property (3) with respect to the history of $\{\theta_t\}$, that is,

$$p(z_t|\theta_t, \theta_{t-1}, \theta_{t-2}, \ldots, \theta_0) = p(z_t|\theta_t). \tag{4}$$

- The process $\{\theta_t\}$ depends on the past observations only through its own history, that is,

$$p(\theta_t|\theta_{t-1}, z_{t-1}, z_{t-2}, \ldots, z_1) = p(\theta_t|\theta_{t-1}). \tag{5}$$

An evolution-observation pair can be illustrated with the following dependency scheme:

$$
\begin{array}{ccccccccc}
\theta_0 & \longrightarrow & \theta_1 & \longrightarrow & \theta_2 & \longrightarrow & \ldots & \longrightarrow & \theta_t & \longrightarrow & \ldots \\
& & \downarrow & & \downarrow & & & & \downarrow & & \\
& & z_1 & & z_2 & & & & z_t & &
\end{array}
$$

Notice, that as soon as a state-space model is defined for an evolution-observation pair, then the assumptions of the model come in parallel with the above definitions (Kaipio & Somersalo, 2005). Assume that the stochastic processes $\{\theta_t\}, \{z_t\}$ form an evolution-observation pair. Then the following problems are under consideration:

- *Prediction*, that is, the determination of $p(\theta_t|z_{t-1}, z_{t-2}, \ldots, z_1)$.

- *Filtering*, that is, the determination of $p(\theta_t|z_t, z_{t-1}, \ldots, z_1)$.

- *Fixed interval smoothing*, that is, the determination of $p(\theta_t|z_T, \ldots, z_t, \ldots, z_1)$, when a complete measurement sequence is available for $t = 1, 2, \ldots, T$.

Based on the conditional or posterior densities, estimators for the states can be defined in a Bayesian framework. It can also be noticed, that all the above problems are computationally related to the prediction problem as an intermediate step.

4. Dynamical estimation of EPs with a linear state-space model

The sampled potential (from channel l) relative to the successive stimulus or trial t can be denoted with a column vector of length M, i.e. $z_t = (z_t(1), z_t(2), \ldots, z_t(M))^T$ for $t = 1, \ldots, T$, where T is the total number of trials, and $(\cdot)^T$ denotes transposition.

A widely used model for EP estimation is the additive noise model (Karjalainen et al., 1999), that is,

$$z_t = s_t + v_t. \tag{6}$$

The vector s_t corresponds to the part of the activity that is related to the stimulation, and the rest of the activity v_t is usually assumed to be independent of the stimulus. Single-trial EPs can be modeled as a linear combination of some pre-selected basis vectors. Then the model takes the form

$$z_t = H_t\theta_t + v_t, \tag{7}$$

where H_t is the observation matrix, which contains the basis vectors $\psi_{t,1}, \ldots, \psi_{t,k}$ of length M in its columns, and θ_t is a parameter vector of length k. The estimated EPs $\hat{s}_t$ can be obtained by using the estimated parameters $\hat{\theta}_t$ as follows:

$$\hat{s}_t = H_t\hat{\theta}_t. \tag{8}$$

The measurement vectors z_t can be considered as realizations of a stochastic vector process, that depend on some unobserved parameters θ_t (state vectors) through (7). For the time evolution of the hidden process θ_t a linear first order Markov model can be used (Georgiadis et al., 2005b), that is,

$$\theta_t = F_t\theta_{t-1} + \omega_t, \tag{9}$$

with some initial distribution for θ_0. Equations (7), (9) form a linear state-space model, where F_t, H_t are preselected matrices. Other assumptions of the model are that for every $i \neq j$ the observation noise vectors v_i, v_j and the state noise vectors ω_i, ω_j are mutually independent and independent of θ_0.

5. Kalman filter and smoother algorithms

Kalman filtering problem is related to the determination of the mean square estimator $\hat{\theta}_t$ for θ_t given the observations $z_1, \ldots, z_t$ (Kalman, 1960). This is equal to the conditional mean

$$\hat{\theta}_t = E\{\theta_t | z_1, \ldots, z_t\} = E\{\theta_t | Z_t\}. \tag{10}$$

The optimal linear mean square estimator can be obtained recursively by restricting to a linear conditional mean, or by assuming v_t and ω_t to be Gaussian (Sorenson, 1980). The Kalman filter algorithm can be written as follows:

- Initialization

$$
\begin{aligned}
C_{\tilde{\theta}_0} &= C_{\theta_0} \tag{11}\\
\hat{\theta}_0 &= E\{\theta_0\} \tag{12}
\end{aligned}
$$

- Prediction step

$$
\begin{aligned}
\hat{\theta}_{t|t-1} &= F_t\hat{\theta}_{t-1} \tag{13}\\
C_{\tilde{\theta}_{t|t-1}} &= F_t C_{\tilde{\theta}_{t-1}} F_t^T + C_{\omega_t} \tag{14}
\end{aligned}
$$

- Filtering step

$$
\begin{aligned}
K_t &= C_{\tilde{\theta}_{t|t-1}} H_t^T (H_t C_{\tilde{\theta}_{t|t-1}} H_t^T + C_{v_t})^{-1} \tag{15}\\
\hat{\theta}_t &= \hat{\theta}_{t|t-1} + K_t(z_t - H_t\hat{\theta}_{t|t-1}) \tag{16}\\
C_{\tilde{\theta}_t} &= (I - K_t H_t) C_{\tilde{\theta}_{t|t-1}}, \tag{17}
\end{aligned}
$$

for $t = 1, \ldots, T$. The matrix K_t is the Kalman gain, $\hat{\theta}_{t|t-1}$ is the prediction of θ_t based on $\hat{\theta}_{t-1}$, and $\hat{\theta}_{t-1} = E\{\theta_{t-1}|z_{t-1}, \ldots, z_1\}$ is the optimal estimate at time $t - 1$.

If all the measurements $z_t, t = 1, \ldots, T$ are available, then the fixed interval smoothing problem can be considered, that is,

$$\hat{\theta}_t^s = E\{\theta_t|z_1, \ldots, z_T\} = E\{\theta_t|Z_T\}. \tag{18}$$

The forward-backward method for the smoothing problem (Rauch et al., 1965), which gives the smoother estimates as corrections of the filter estimates is complete through the backward recursion:

- Smoothing

$$A_t = C_{\tilde{\theta}_t} F_{t+1}^T C_{\tilde{\theta}_{t+1|t}}^{-1} \tag{19}$$

$$\hat{\theta}_t^s = \hat{\theta}_t + A_t(\hat{\theta}_{t+1}^s - \hat{\theta}_{t+1|t}) \tag{20}$$

$$C_{\tilde{\theta}_t^s} = C_{\tilde{\theta}_t} + A_t(C_{\tilde{\theta}_{t+1}^s} - C_{\tilde{\theta}_{t+1|t}})A_t^T, \tag{21}$$

for $t = T - 1, T - 2, \ldots, 1$. For initialization of the backward recursion the filter estimates are used, i.e. $\hat{\theta}_T^s = \hat{\theta}_T$.

6. EP estimation based on a generic model

The following state-space model for dynamical estimation of evoked potentials is here considered:

$$\theta_t = \theta_{t-1} + \omega_t \tag{22}$$

$$z_t = H\theta_t + v_t, \tag{23}$$

with the selections $F_t = I, t = 1, \ldots, T$, i.e. a random walk model, and $H_t = H$ for all t.

The observation model can be formed from the impulse response of an FIR filter. Consider a linear (non-causal) finite response filter with impulse function defined by the sequence $\{h(n)\}$ over the interval $-M \leq n \leq M$. For a given input $z_t(n), n = 1, \ldots, M$ the output is given by

$$y_t(n) = \sum_{k=-\infty}^{\infty} h(n-k)z_t(k) = \sum_{k=1}^{M} h(n-k)z_t(k), \tag{24}$$

where $z_t(n) = 0$ for $n < 1$.

The output of the filter $y_t = (y_t(1), y_t(2), \ldots, y_t(n), \ldots, y_t(M))^T$ in terms of the input vector $z_t = (z_t(1), z_t(2), \ldots, z_t(n), \ldots, z_t(M))^T$, for $n = 1, \ldots, M$, is given in a compact matrix form by

$$y_t = \begin{bmatrix} h(0) & h(-1) & \ldots & h(1-M) \\ h(1) & h(0) & \ldots & h(2-M) \\ \vdots & \vdots & \vdots & \vdots \\ h(n-1) & h(n-2) & \ldots & h(n-M) \\ \vdots & \vdots & \vdots & \vdots \\ h(M-1) & h(M-2) & \ldots & h(0) \end{bmatrix} z_t, \tag{25}$$

where the filter operator P, i.e. $y_t = Pz_t$, contains time-shifted versions of the impulse function in its columns. The performance of the filter can be approximated by choosing less vectors to form an observation model H with k columns, selected for $i = 1, \ldots, k$ as

$$\psi_i = (h(-d_i), \ldots, h(M - 1 - d_i))^T, \qquad (26)$$

where d_i can be selected based on the values $0, M/(k-1), 2M/(k-1), \ldots, M$. An approximation of the filter performance can be obtained, for example, through the matrix $H(H^T H)^{-1} H^T$ in the ordinary least squares sense. Different observation models, for example, the Gaussian basis (Georgiadis et al., 2005b; Qiu et al., 2006; Ranta-aho et al., 2003), here seen as a low pass filter, can be used.

For the covariances of the state and observation noise processes the choices $C_{\omega_t} = \sigma_\omega^2 I$, $C_{v_t} = \sigma_v^2 I$ for every trial t can be made. Then, the selection of the last variance term is not essential, since only the ratio $\sigma_v^2 / \sigma_\omega^2$ has effect on the estimates. A detailed proof can be found in (Georgiadis et al., 2007). Then the choice $C_{v_t} = I$ can be made, and care should be given to the selection of only one parameter σ_ω^2. In general, if it is tuned too small fast fluctuation of EPs are going to be lost, and if it is selected too big the estimates have too much variance. The selection can be based on experience and visual inspection of the estimates as a balance between preserving expected dynamic variability and greater noise reduction. Extensive discussion and examples related to the selection of this parameter can be found in (Georgiadis, 2007; Georgiadis et al., 2005b; 2007).

7. Examples

7.1 Amplitude variability

In this example, measurements were obtained from an EP experiment with visual stimulation. 310 fixed intensity flash stimuli (red squares) were presented to the subject through a monitor (screen 36.5 x 27.6 cm, distance 1 m). The stimuli were randomly presented every 1.5s (from 1.3s to 1.7s) and their duration was 0.3s. The measurement device was BrainAmp MR plus and the sampling rate was $Fs = 5000$Hz. Prior to the estimation procedure the EEG channels were band pass filtered with pass band 1-500Hz. Then epochs of 0.5s relative to the presentation of stimuli were sampled from channel Oz. All the epochs were kept for estimation.

The observation model was created based on a low pass FIR filter with impulse response obtained by truncating an ideal low pass filter (sinc function) with a Hanning window. The cut-off frequency was selected to be $f_c = 20$Hz and the number of vectors was selected to be $k = 21$. The empirical rule:

$$k = \left[\frac{f_c}{F_s/2} M \right] + 1, \qquad (27)$$

where $[\cdot]$ denotes integer part, seemed to produce good values for k for different values of F_s, f_c, M. The selected observation model is illustrated in Figure 3, where the columns of the matrix H are represented as rows in an image plot.

Kalman filter and smoother estimates were computed for the model (22), (23) with the selection $\sigma_\omega^2 = 1$. The value was chosen empirically by visual examination of the estimates. For initialization of the algorithms, half of the measurements were used in a backward recursion with Kalman filter algorithm. The last (converged) estimates were used to initialize the Kalman filter forward run. For the initialization of the final backward recursion (Kalman smoother) the filter estimates were used.
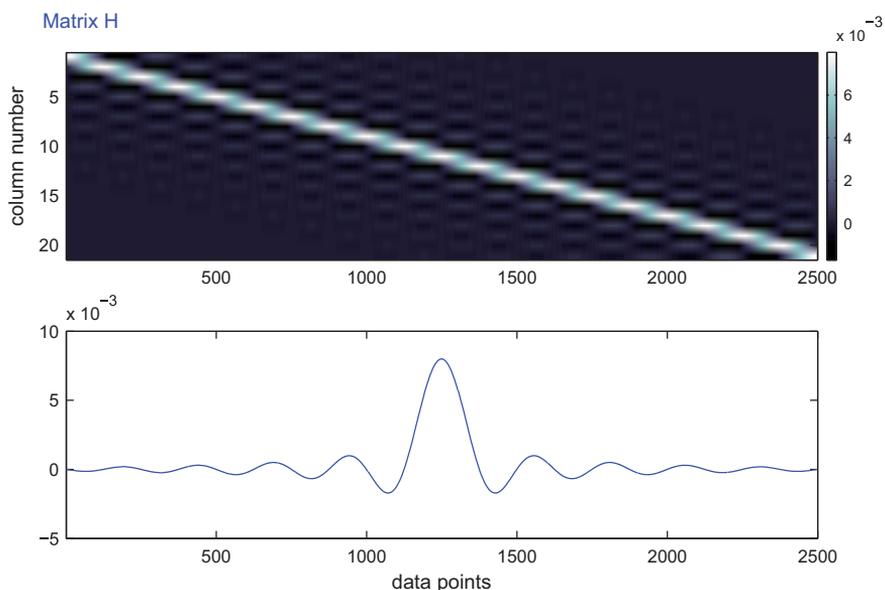
Fig. 3. The selected observation model. Up: the columns of the matrix *H* as rows in an image plot. Down: the 11th column.

Figure 4 (top, left) shows the noisy EP measurements as an image plot. The positive dominant peak, here occurring about 160 ms after visual stimulation, is visible at the center of the image. The obtained estimates are presented in the same figure for Kalman filter (top, right) and smoother (bottom left). The averaged EPs obtained from the raw measurements and from the estimates are also seen in the middle of the figure. The positive dominant peak can be observed from this plot. Clearly, the time variation of the EPs is revealed. A decrease in amplitude of the dominant positive peak is clearly observable, suggesting possible habituation to the stimuli presentation. The amplitude of the peak, estimated simply as the maximum value within the time interval 100-200ms after the presentation of the stimuli, is also plotted as a function of the successive stimulus *t*. Furthermore, the time-varying latency of the peak is presented. From these plots it can be easier observed the gradual decrease of the amplitude. Finally, the improvement due to the smoothing procedure is visible. The smoother algorithm cancels the time-lag of the filtering procedure. In parallel, it removes greater the noise, thus improving the latency estimation, especially for the very weak evoked potentials.

7.2 Latency variability
In this example, measurements related to the P300 event related potential were used. The P300 peak is one of the most extensively studied cognitive potential and there exist many studies where the trial-to-trial variability of the component is discussed, for example, (Holm et al., 2006).
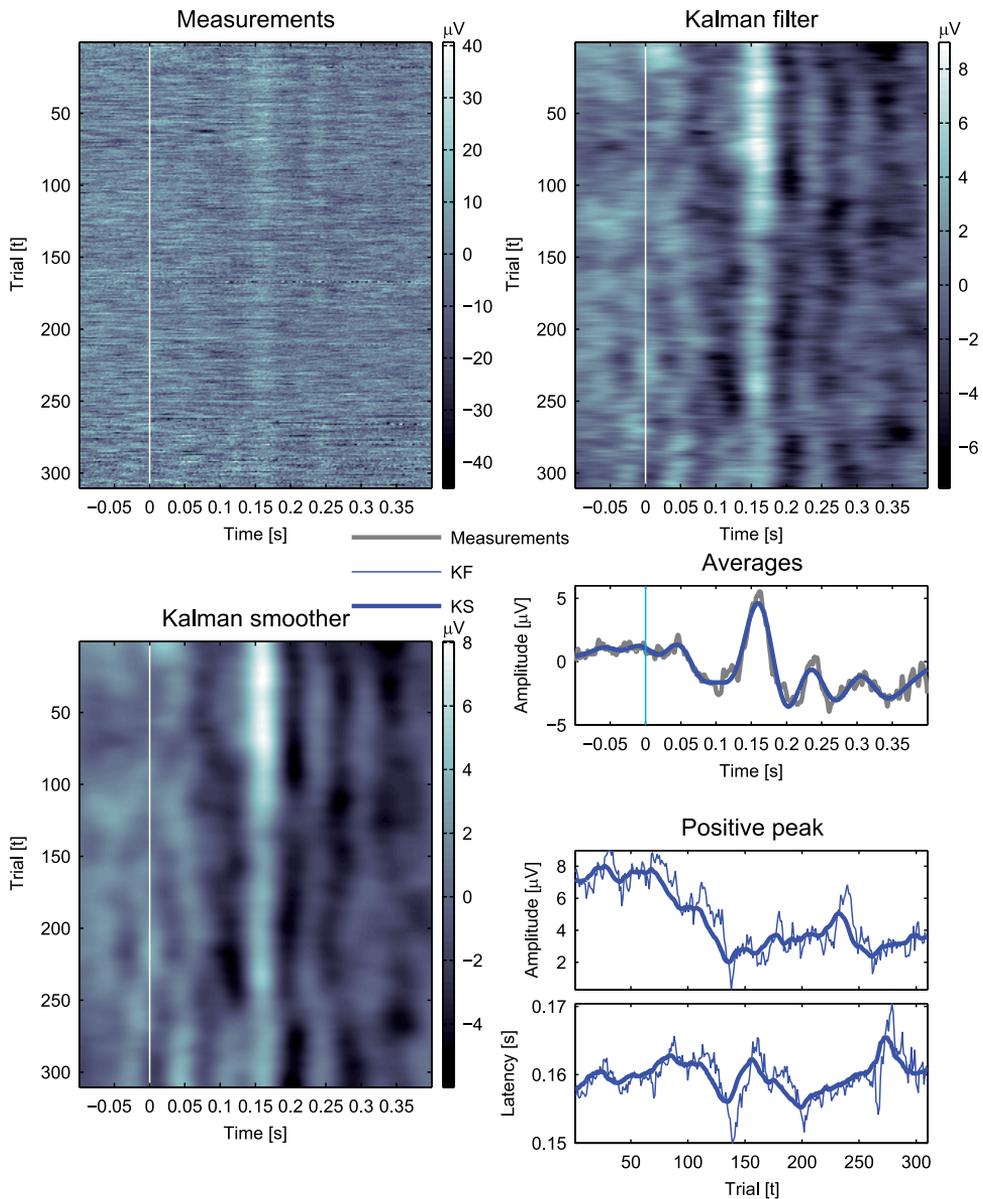
Visual stimulation (channel Oz)

Fig. 4. Single-trial EP amplitude variability.

EEG measurements were obtained from a standard oddball paradigm with auditory stimulation. During the recording, 569 auditory stimuli were presented with an inter-stimulus inter-

val of 1s, 85% of the stimuli at 800Hz and randomly presented 15% deviant tones at 560Hz. The subject was sitting in a chair and was asked to press a button every time he heard the deviant target tone. The sampling rate of the EEG was 500 Hz. From the recordings, channel Cz was selected for analysis, after bandpass filtering in the range 1-40Hz. Average responses from the two conditions are shown in Figure 2 (Section 2). For investigation of the single trial variability of the P300 peak, EEG epochs from -100 ms to 600 ms relative to the stimulus onset of each deviant stimulus were here used.

The model was designed as in section 7.1 but now for the slower P300 wave the selection $f_c = 10$Hz was made. The application of the empirical rule (27) gave in this case $k = 15$. Kalman smoother estimates were computed with the selection $\sigma_\omega^2 = 9$, with respect to the expected faster variability of the potential.

In Figure 5 (I) there are presented the EP measurements in the original stimulus order (trial-by-trial). In the same figure (II) the obtained estimates based on the measurements (I) are shown. Clearly, in the estimates, the dynamic variability of the P300 peak potential is revealed, suggesting that it cannot be considered as occurring at fixed latency from the stimuli presentation. At the same image (II), the estimated latency is also plotted as a function of the consecutive trial t. The latency of the peak was estimated from the Kalman smoother estimates based on the maximum value within the time interval 250-370ms after the presentation of the stimuli.

The estimated time-varying latency of the P300 peak was then used to order the single-trial measurements. The sorted single-trials (condition-by-condition) are shown at Figure 5 (III). The shorted latency estimates are plotted again over the image plot. This plot clearly demonstrates that the latency estimates obtained with Kalman smoother are of acceptable accuracy.

Finally, the algorithm was also applied to the sorted measurements (III). The value $\sigma_\omega^2 = 4$ was selected and new point estimates for the latency were obtained as before. Kalman smoother estimates and the new latency estimates are plotted in Figure 5 (IV). The linear trend of the sorted potentials allows the use of even smaller value for state-noise variance parameter (Georgiadis et al., 2005b), thus reducing even more the noise without reducing the variability of the peak. The last obtained estimates of the latencies were plotted over the original non sorted measurements (I). The similarities between the estimated latency fluctuations in (I) and (II) underline the robustness of the method.

8. Conclusion and Future Directions

EP research has to deal with several inherent difficulties. Traditional analysis is based on averaged data often by forming extra grand averages of different populations. Thus, trial-to-trial variability and individual subject characteristics are largely ignored (Fell, 2007). Therefore, the study of isolated components retrieved by averages might be misleading, or at least it is a simplification of the reality. For example, habituation may occur and the responses could be different from the beginning to the end of the recording session. Furthermore, cognitive potentials exhibit rich latency and amplitude variability that traditional research based on averaging is not able to exploit for studying complex cognitive processes. Latency variability could be used, for instance, for studying perceptual changes, quantifying stimulus classification speed or task difficulty.

In this chapter, state-space modeling for single-trial estimation of EPs was presented in its general form based on Bayesian estimation theory. This formulation enables the selection of different models for dynamical estimation. In general, the applicability of the proposed
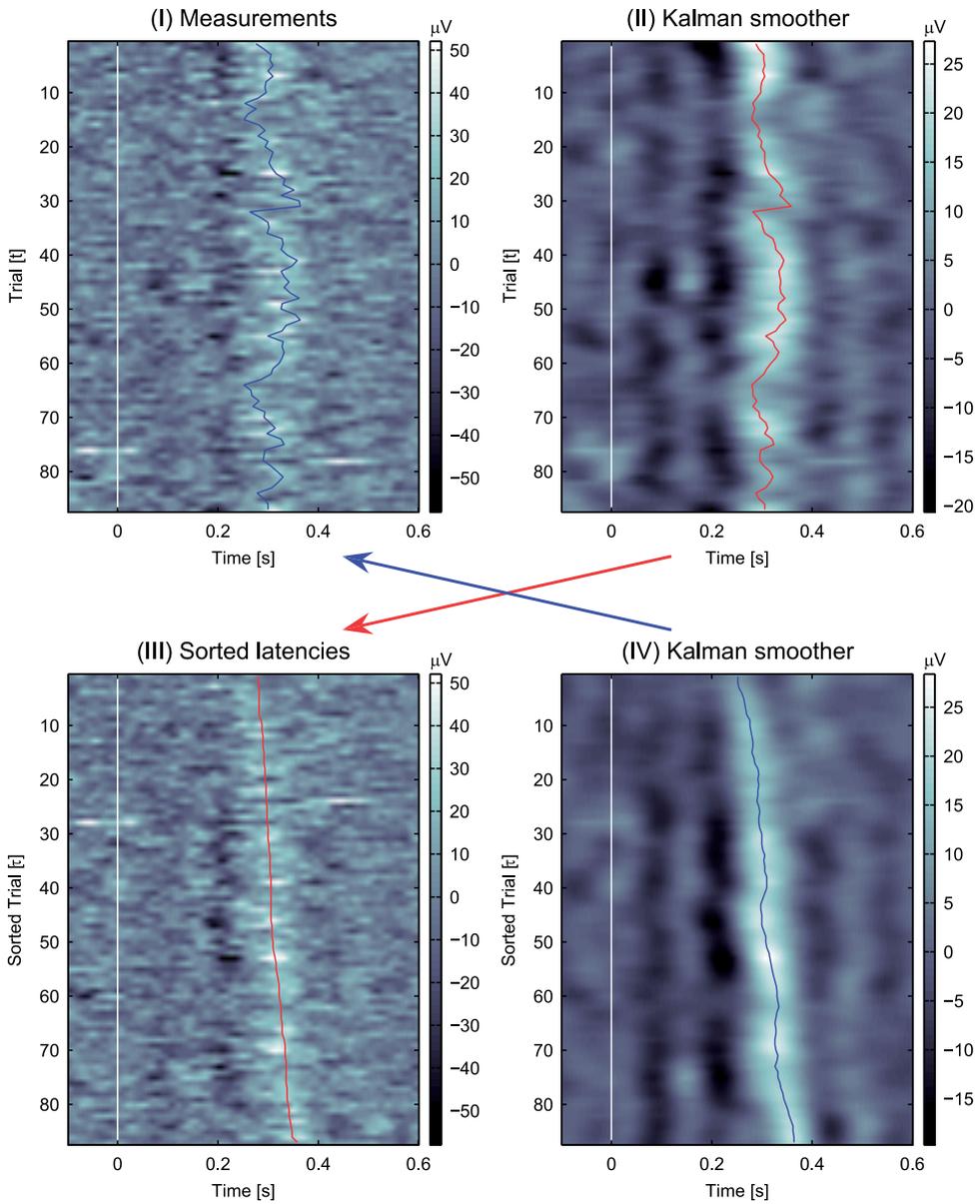
Auditory stimulation (channel Cz)

Fig. 5. Single-trial EP latency variability.

methodology primarily relates on the assumption of hidden dynamic variability from trial-to-trial or from condition-to-condition. A practical method for designing an observation model was also presented and its capability to reveal meaningful amplitude and latency fluctuations in EP measurements was demonstrated. In the approach, optimal estimates for the states are obtained with Kalman filter and smoother algorithms. When all the measurements are available (batch processing) Kalman smoother should be used.

EPs also contain rich spatial information that can be used for describing brain dynamics (Makeig et al., 2004; Ranta-aho et al., 2003). In this study, this important issue was not discussed and emphasis was given on optimal estimation of some temporal EP characteristics. Future development of the presented methodology involves the extension of the approach to multichannel and multimodal data sets, for instance, simultaneously measured EEG/ERP and fMRI/BOLD signals (Debener et al., 2006), for the study of dynamic changes of the central nervous system.

Acknowledgments

9. References

Cerutti, S., Bersani, V., Carrara, A. & Liberati, D. (1987). Analysis of visual evoked potentials through Wiener filtering applied to a small number of sweeps, *Journal of Biomedical Engineering* **9**(1): 3–12.

Debener, S., Ullsperger, M., Siegel, M. & Engel, A. (2006). Single-trial EEG-fMRI reveals the dynamics of cognitive function, *Trends in Cognitive Sciences* **10**(2): 558–63.

Delorme, A. & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis, *Journal of Neuroscience Methods* **134**(1): 9–21.

Doncarli, C., Goering, L. & Guiheneuc, P. (1992). Adaptive smoothing of evoked potentials, *Signal Processing* **28**(1): 63–76.

Fell, J. (2007). Cognitive neurophysiology: Beyond averaging, *NeuroImage* **37**: 1069–1027.

Georgiadis, S. (2007). *State-Space Modeling and Bayesian Methods for Evoked Potential Estimation*, PhD thesis, Kuopio University Publications C. Natural and Environmental Sciences 213. (available: http://bsamig.uku.fi/).

Georgiadis, S., Ranta-aho, P., Tarvainen, M. & Karjalainen, P. (2005a). Recursive mean square estimators for single-trial event related potentials, *Proc. Finnish Signal Processing Symposium - FINSIG'05*, Kuopio, Finland.

Georgiadis, S., Ranta-aho, P., Tarvainen, M. & Karjalainen, P. (2005b). Single-trial dynamical estimation of event related potentials: a Kalman filter based approach, *IEEE Transactions on Biomedical Engineering* **52**(8): 1397–1406.

Georgiadis, S., Ranta-aho, P., Tarvainen, M. & Karjalainen, P. (2007). A subspace method for dynamical estimation of evoked potentials, *Computational Intelligence and Neuroscience* **2007**: Article ID 61916, 11 pages.

Georgiadis, S., Ranta-aho, P., Tarvainen, M. & Karjalainen, P. (2008). Tracking single-trial evoked potential changes with Kalman filtering and smoothing, *30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Vancouver, Canada, pp. 157–160.

Holm, A., Ranta-aho, P., Sallinen, M., Karjalainen, P. & Müller, K. (2006). Relationship of P300 single trial responses with reaction time and preceding stimulus sequence, *International Journal of Psychophysiology* **61**(2): 244–252.

Intriligator, J. & Polich, J. (1994). On the relationship between background EEG and the P300 event-related potential, *Biological Psychology* **37**(3): 207–218.

Jansen, B., Agarwal, G., Hegde, A. & Boutros, N. (2003). Phase synchronization of the ongoing EEG and auditory EP generation, *Clinical Neurophysiology* **114**(1): 79–85.

Kaipio, J. & Somersalo, E. (2005). *Statistical and Computational Inverse Problems*, Applied Mathematical Sciences, Springer.

Kalman, R. (1960). A new approach to linear filtering and prediction problems, *Transactions of the ASME, Journal of Basic Engineering* **82**: 35–45.

Karjalainen, P., Kaipio, J., Koistinen, A. & Vauhkonen, M. (1999). Subspace regularization method for the single trial estimation of evoked potentials, *IEEE Transactions on Biomedical Engineering* **46**(7): 849–860.

Knuth, K., Shah, A., Truccolo, W., Ding, M., Bressler, S. & Schroeder, C. (2006). Differentially variable component analysis (dVCA): Identifying multiple evoked components using trial-to-trial variability, *Journal of Neurophysiology* **95**(5): 3257–3276.

Li, R., Principe, J., Bradley, M. & Ferrari, V. (2009). A spatiotemporal filtering methodology for single-trial ERP component estimation, *IEEE Transactions on Biomedical Engineering* **56**(1): 83–92.

Makeig, S., Debener, S. & Delorme, A. (2004). Mining event-related brain dynamics, *Trends in Cognitive Science* **8**(5): 204–210.

Makeig, S., Westerfield, M., Jung, T.-P., Enghoff, S., Townsend, J., Courchesne, E. & Sejnowski, T. (2002). Dynamic brain sources of visual evoked responses, *Science* **295**: 690–694.

Mäkinen, V., Tiitinen, H. & May, P. (2005). Auditory even-related responses are generated independently of ongoing brain activity, *NeuroImage* **24**(4): 961–968.

Malmivuo, J. & Plonsey, R. (1995). *Bioelectromagnetism*, Oxford university press, New York.

Niedermeyer, E. & da Silva, F. L. (eds) (1999). *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, 4th edn, Williams and Wilkins.

Qiu, W., Chang, C., Lie, W., Poon, P., Lam, F., Hamernik, R., Wei, G. & Chan, F. (2006). Real-time data-reusing adaptive learning of a radial basis function network for tracking evoked potentials, *IEEE Transanctions on Biomedical Engineering* **53**(2): 226–237.

Quiroga, R. Q. & Garcia, H. (2003). Single-trial evoked potentials with wavelet denoising, *Clinical Neurophysiology* **114**: 376–390.

Ranta-aho, P., Koistinen, A., Ollikainen, J., Kaipio, J., Partanen, J. & Karjalainen, P. (2003). Single-trial estimation of multichannel evoked-potential measurements, *IEEE Transactions on Biomedical Engineering* **50**(2): 189–196.

Rauch, H., Tung, F. & Striebel, C. (1965). Maximum likelihood estimates of linear dynamic systems, *AIAA Journal* **3**: 1445–1450.

Sorenson, H. (1980). *Parameter Estimation, Principles and Problems*, Vol. 9 of *Control and Systems Theory*, Marcel Dekker Inc., New York.

Thakor, N., Vaz, C., McPherson, R. & Hanley, D. F. (1991). Adaptive Fourier series modeling of time-varying evoked potentials: Study of human somatosensory evoked response to etomidate anesthetic, *Electroencephalography and Clinical Neurophysiology* **80**(2): 108–118.

Truccolo, W., Mingzhou, D., Knuth, K., Nakamura, R. & Bressler, S. (2002). Trial-to-trial vari-
    ability of cortical evoked responses: implications for the analysis of functional con-
    nectivity, *Clinical Neurophysiology* **113**(2): 206–226.
Turetsky, B., Raz, J. & Fein, G. (1989). Estimation of trial-to-trial variation in evoked potential
    signals by smoothing across trials, *Psychophysiology* **26**(6): 700–712.

New Developments in Biomedical Engineering

Edited by Domenico Campolo

Biomedical Engineering is a highly interdisciplinary and well established discipline spanning across engineering, medicine and biology. A single definition of Biomedical Engineering is hardly unanimously accepted but it is often easier to identify what activities are included in it. This volume collects works on recent advances in Biomedical Engineering and provides a bird-view on a very broad field, ranging from purely theoretical frameworks to clinical applications and from diagnosis to treatment.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Stefanos Georgiadis, Perttu Ranta-aho, Mika Tarvainen and Pasi Karjalainen (2010). State-Space Modeling for Single-Trial Evoked Potential Estimation, New Developments in Biomedical Engineering, Domenico Campolo (Ed.), ISBN: 978-953-7619-57-2, InTech, Available from: http://www.intechopen.com/books/new-developments-in-biomedical-engineering/state-space-modeling-for-single-trial-evoked-potential-estimation

INTECH
open science | open minds