Subspace Regularization method for the Single Trial Estimation of Multi Channel Evoked Potential Measurements

P.O. Ranta-aho , A.S. Koistinen, J.O. Ollikainen, J.P. Kaipio, J. Partanen, P.A. Karjalainen

Abstract— **A method for single trial estimation of multi channel evoked potentials is presented. The proposed method is based on the regularized least squares scheme. The spatial correlation between the channels is used as additional information in the estimation procedure. Amplitude estimates obtained with proposed method is compared to the estimates calculated without using the spatial information. The performance of the method is evaluated using simulated and real data of P300 responses measured using auditory stimuli. The multi channel approach is shown to give realistic and comparable information about the amplitude differences of the P300 peak between different channels.**

Keywords— **Evoked potentials, single trial estimation, multi channel measurements, regularized least squares**

I. Introduction

Event related potentials (ERP) have been widely used for studying brain activity associated with higher mental functions. The most common way to measure the parameters of ERPs is to take an average over time locked single trial measurements. The implicit assumption in the averaging is that the task-related cognitive process does not vary much in timing from trial to trial. However, it has been evident for few decades that in many cases this assumption is not valid [1]. The observation of the variation of the parameters of the ERPs permits the dynamical assessment of the changes in the cognitive state of human. Thus the goal in the analysis of the ERPs is currently the estimation of the single potentials, a task, that we call single trial estimation. With the single trial ERPs it is possible to e.g. examine the habituation during the test or the level of attention.

Common approach to denoise single trial ERPs is to form a filter which will filter out the unwanted contribution of the on-going background activity of the brain, as well as possible. However, the problem in this task is often very low signal-to-noise ratio. Basically filtering is an estimation problem in which the measurement is used as data in estimation of the assumed underlying signal. One possibility to improve the estimate is to apply additional information about the underlying signal, i.e. the evoked potentials, to the estimation. This information can be concerned with the assumed smoothness of the evoked potentials or we may use some limits for the possible locations of the peaks in the potentials. This kind of assumptions are employed e.g. in the time varying filter introduced in [2]. The implementation of the method, however, necessitates the knowledge of the covariance of the evoked potentials. That is in fact another estimation problem that has no easy solution. An interesting way to add temporal information to the estimation is to use the so-called regularized least squares method, which as other regularization methods has its origin in the theory of ill-posed inverse problems [3]. The method has previously been used in [4] in case of single channel measurements. The main benefit of the method is its easy implementation for different kinds of data.

In practice the evoked potential measurements are usually performed with multiple electrodes. This means that also spatial information is gathered in the measurements. Obviously we should not neglect this information in the estimation of the single trial ERPs. One possible solution is to interpret the spatial correlation of the multi channel measurements in the form of additional information. This type of multi channel time varying Wiener filter is introduced in [5]. The proposed filter is optimal in the mean square sense in the class of unbiased estimators. However, the implementation of the method is quite complicated and necessitates basically information that is not available prior to the estimation. Another approach to take the additional information of the spatial correlation of the channels into account is again to use the regularization approach. The approach is briefly introduced in [6].

In this paper we introduce a systematic method for single trial multi channel measurements. The method is based on the regularized least squares scheme and it uses both spatial and temporal information in the estimation. Introduced method is extension of the single channel method presented in [4]. These methods are also compared to each other using simulations and the real measurements. The results show that the multi channel approach preserves the amplitude information of the peaks of the estimated potential better than the single channel approach.

The rest of the paper is organized as follows. The mathematical basis of the subspace regularization method as well as the single channel method for single trial ERP estimation is reviewed in Section II. An extension of the method to multi channel measurements is derived in Section III. The choice of the regularization parameters is shortly discussed in Section III-A. Properties of the background EEG can also be taken into account in the esti-

P.O. Ranta-aho, P.A. Karjalainen, J. Ollikainen and J.Kaipio, are with University of Kuopio, Department of Applied Physics, P.O.Box 1627, FIN-70211 Kuopio, Finland

A.S. Koistinen is with Finnish Institute of Occupational Health, Brain@Work Laboratory, Topeliuksenkatu 41 a A, FIN-00250 Helsinki, Finland

J. Partanen is with Kuopio University Hospital, Department of Clinical Neurophysiology, P.O.Box 1777, FIN-70211 Kuopio, Finland

mation with proposed methods, this is described briefly in Section III-B. Systematic procedure for applying the multi channel method is described in Section IV. In Section V methods are compared using both realistic simulated ERPs, Section V-B, and real measurement data Section V-C. The applicability of the methods is discussed in Section VI.

II. Estimation with linear observation model

A widely used model for evoked potentials is the linear additive noise model, in which the observations are of the form

$$z_i = s_i + e_i. \tag{1}$$

We denote here the sampled potential measurement after i'th stimulus with a column vector z_i. The length of the vector z_i is T. The evoked potential s_i corresponds to the part of the activity that is correlated with the stimulus. It is usually a transient waveform that consists of activity peaks of some duration. The other part of activity, the background EEG e_i, is usually thought to be independent of the stimulus and the evoked potential s_i. Thus the electrical activity of the brain is the source for both s_i and e_i.

The evoked potential s_i can be further modeled as a linear combination of some pre-selected basis vectors ψ_j. The vectors s_i can then be written in the form

$$s_i = H\theta_i \tag{2}$$

Where H is a matrix that contain the basis vectors $\psi_1, \ldots, \psi_k$ in its columns and θ_i is a length k vector of parameters. Thus the observation model is now

$$z_i = H\theta_i + e_i \tag{3}$$

In the modeling of the evoked potentials we then have to estimate parameters θ_i based on the measured data z_i using some estimation criterion. The estimated evoked potentials $\hat{s}_i$ can then be obtained using the estimated parameters $\hat{\theta}_i$ with equation

$$\hat{s}_i = H\hat{\theta}_i \tag{4}$$

A. Selection of the basis

With linear observation model the choice of the observation matrix H has a significant role. The best choice would be the true physical model which is capable to model both temporal and spatial information in the measurements. However, that necessitates the modeling of the sources as physical current generators as well as modeling the electrical properties of the head. This is obviously not a trivial task. In this work we use a simplier phenomenological model.

First we assume that the potentials consist of positive and negative humps. Sampled Gaussian or sigmoid functions can then be a good choice for the basis. We call this kind of basis that do not depend on the data a generic basis. Other choices could be e.g. Fourier basis or several possibilities of wavelet bases. It is also possible to use the p first eigenvectors of the correlation matrix of the measurements as a basis. The least squares solution with this basis is then equivalent to the so-called principal component regression approach [7].

Both the use of generic basis vectors and the principal component regression approach have their own benefits. On one hand the generic vectors are usually robust and easy to generate. They also commonly model different intervals of the measurement vector in a homogeneous way. On the other hand the eigenvector basis is optimal for a set of measurements, although this is strictly true only for jointly Gaussian measurements. These two approaches can be combined together, as we will show in the following sections.

B. Regularized least squares

The task is to estimate the parameters θ_i in (4) based on the measurements z_i. We state a generalized least squares solution for the parameters θ_i

$$\hat{\theta}_i = \arg\min_{\theta_i} \left\{ \|(z_i - H\theta_i)\|^2 + \alpha^2 \|L\theta_i\|^2 \right\}. \tag{5}$$

The statement arg means the argument of the expression. The solution of (5) is called the generalized Tikhonov regularized solution [8] and it is clearly a modification of the ordinary least squares solution $\hat{\theta}_i^{\mathrm{LS}} = \arg\min_{\theta_i} \left\{ \|(z_i - H\theta_i)\|^2 \right\}$ to the direction in which the semi norm $\|L\theta_i\|$, the so-called *side constraint*, gets smaller.

The so-called regularization matrix L in the side constraint can be e.g. the second derivative approximation [9]. Minimization of the second derivative obviously smoothes sharp spikes in the estimated vector when compared to the ordinary least squares solution.

It is easy to show [10] that the regularized solution can be written in the form

$$\hat{\theta}_i = (H^T H + \alpha^2 L^T L)^{-1} H^T z_i \tag{6}$$

C. Subspace regularization method

The principal component approach can be combined with the generic basis vectors using the regularized least squares method as we will show here. We use a generic Gaussian basis vectors in the columns of the observation matrix H. Next we calculate the eigenvectors of the correlation matrix R_z of the measurements and use p first eigenvectors as the columns of the matrix $H_{\mathcal{S}}$. The matrix $H_{\mathcal{S}}$ will now contain an orthonormal basis of the subspace $\mathcal{S}$. We want that the estimated evoked potential is close to this subspace. The projection of $s_i = H\theta_i$ onto $\mathcal{S}$ is $(H_{\mathcal{S}} H_{\mathcal{S}}^T) H\theta_i$ and the distance of s_i from $\mathcal{S}$ can be written in the form $\left\|(I - H_{\mathcal{S}} H_{\mathcal{S}}^T) H\theta_i\right\|$. Remembering that we should construct such a matrix L that the side constraint $\|L\theta_i\|$ in (5) is small for all expectable θ_i, we thus select $L = (I - H_{\mathcal{S}} H_{\mathcal{S}}^T) H$. Since $L^T L = H^T (I - H_{\mathcal{S}} H_{\mathcal{S}}^T) H$, the desired solution for the parameters θ_i can be written in the form

$$\hat{\theta}_i = (H^T H + \alpha^2 H^T (I - H_{\mathcal{S}} H_{\mathcal{S}}^T) H)^{-1} H^T z_i \tag{7}$$

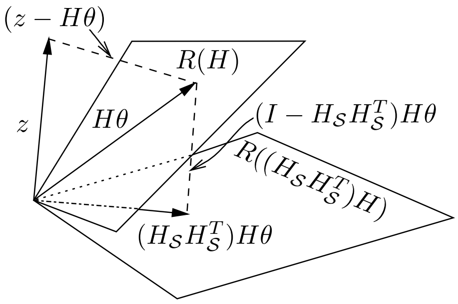

Fig. 1. A graphical interpretation of the subspace regularization.

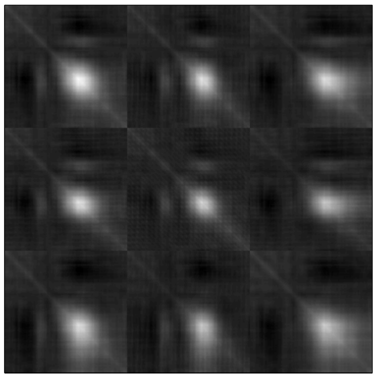

Fig. 2. Correlation matrix R_z of the multi channel measurements in the case of three channels. Light areas correspond large correlation.

The estimate for the evoked potential is then

$$\hat{s}_i = H\hat{\theta}_i \tag{8}$$

This is called the subspace regularized solution [4].

In the subspace regularization method we do not restrict the evoked potentials to be strictly a linear combination of either set of the basis vectors. We rather use one set of vectors as a model for the evoked potentials, these are the columns of the matrix H. Eigenvectors of the covariance matrix of the measurements represent the correlation between the set of measurements. The parameter α controls the weight between the different sets. A graphical interpretation of the subspace regularization is shown in Fig. 1. The Bayesian aspects of the subspace regularization method are discussed in [11].

III. Single trial estimation of multi channel measurements

Although the evoked potential and the background EEG are independent by definition both the activities s_i and e_i are highly correlated between different channels $j = 1, \ldots, M$. This spatial correlation could be taken into account in the formulation of the observation matrix H. However, this task would necessitate the true physical modeling of the head as a volume conductor. We use here much more simplier approach which is based on the subspace regularization method in the following way.

For the multi channel measurements we use the following notation. Let $z_i^{(j)}$ denote the evoked potential for the i'th stimulus measured using the j'th channel. We concatenate the measurements in the following way.

$$z_i = \begin{pmatrix} z_i^{(1)} \\ \vdots \\ z_i^{(M)} \end{pmatrix} = \begin{pmatrix} s_i^{(1)} \\ \vdots \\ s_i^{(M)} \end{pmatrix} + \begin{pmatrix} e_i^{(1)} \\ \vdots \\ e_i^{(M)} \end{pmatrix} \tag{9}$$

The vector z_i is thus a vector of length $M \times T$, where M is the number of channels.

As in the single channel case we use a generic Gaussian basis for each channel separately. The columns of the observation matrix $H^{(j)}$ are thus Gaussian shaped vectors. Modeling the channels separately means that the matrix H is a block diagonal matrix

$$s_i = \begin{pmatrix} H^{(1)} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & H^{(M)} \end{pmatrix} \begin{pmatrix} \theta_i^{(1)} \\ \vdots \\ \theta_i^{(M)} \end{pmatrix}, \tag{10}$$

where the matrices $H^{(j)}$ are identical. This model itself does not contain any dependence between the channels. Next we form the eigendecomposition of the data correlation matrix

$$R_z = \begin{pmatrix} R_z^{(1,1)} & \ldots & R_z^{(1,M)} \\ \vdots & \ddots & \vdots \\ R_z^{(M,1)} & \ldots & R_z^{(M,M)} \end{pmatrix} \tag{11}$$

and use a few eigenvectors corresponding the largest eigenvalues to form the regularization matrix $H_{\mathcal{S}}$. Typical correlation matrix of real measurement data of three channels (FZ, CZ, PZ) is shown in Fig. 2 and in Fig. 3 are shown the eigenvectors of the correlation matrix corresponding the four largest eigenvalues.

Because the correlation is calculated using the stacked measurements z_i, the eigenvectors model also the correlation between the separate channels. As in the single channel case the estimates for the evoked potentials can be obtained with the equations (7) and (4).

A. Choice of the regularization parameter

A common problem with the regularization based methods is that the choice of optimal value of the regularization parameter α can be tricky. Several methods have been developed for the estimation of the optimal value of the parameter. The so-called GCV (Generalized Cross Validation) method [12] is one of the possibilities. The use of GCV in the selection of the regularization parameter in the single channel approach is discussed in [4]. However, the estimation methods for single trial evoked potentials proposed here are quite robust to small changes of α. Thus here our selection of the regularization parameter is based on the experience. Typical values of α are from 5 to 20. If the value of α is too small the estimated evoked potential will be non smooth and follow data closely. Too large value will force the estimation be near constant zero.

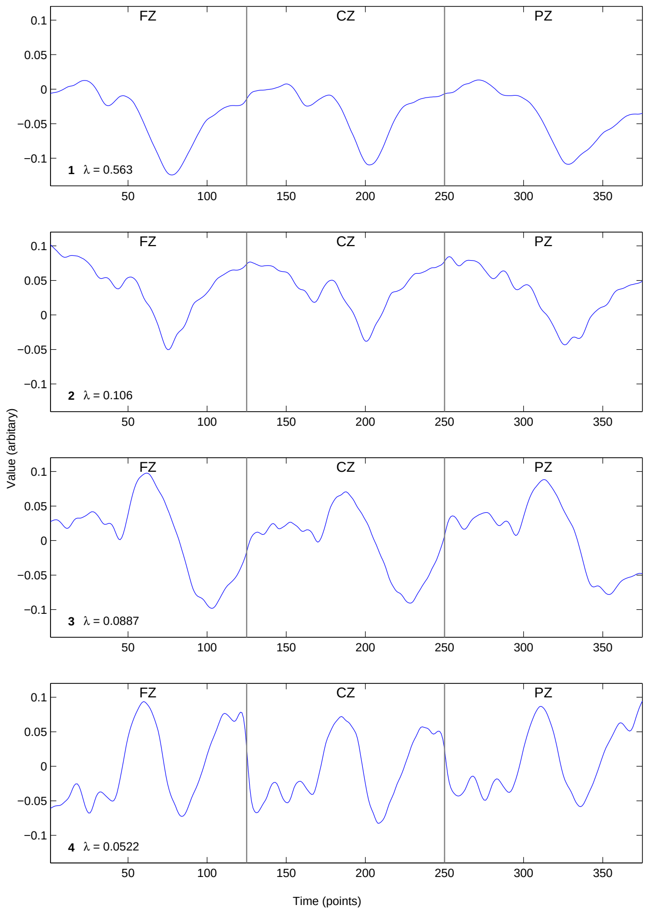

Fig. 3. The eigenvectors corresponding the four largest eigenvalues λ of the correlation matrix of the real measurement data in the multi channel case. The vertical lines show the channel borders.

B. Taking background EEG into account

If the covariance of the background EEG can be estimated, both the single channel and multi channel approaches can be expanded to take the background EEG into account in the estimation. The solution for parameters $\hat{\theta}$ (5) can be written in a form

$$\hat{\theta}_i = \arg\min_{\theta_i} \left\{ \|L_1(z_i - H\theta_i)\|^2 + \alpha^2 \|L\theta_i\|^2 \right\} \qquad (12)$$

which is a modification of the ordinary weighted least squares solution $\hat{\theta}_{\mathrm{LS}} = \arg\min_{\theta_i} \left\{ \|L_1(z_i - H\theta_i)\|^2 \right\}$. The equation (6) will now be

$$\hat{\theta}_i = (H^T L_1^T L_1 H + \alpha^2 L^T L)^{-1} H^T L_1^T L_1 z_i. \qquad (13)$$

Now we can select $L_1^T L_1 = C_v^{-1}$, where C_v is the estimate for the covariance of the background EEG. The estimates for the evoked potentials can again be calculated with equation (4).

Taking the background EEG into account will make the method more sensitive for small changes of regularization parameter α and thus the choice of α will be more crucial. In case studies later we have not used any information of the background EEG i.e. we have chosen $C_v \equiv I$ to avoid the disturbances of the possible non optimal selection of the α. However, the contribution of the background EEG is taken into account in the systematic method presented in the following section.

IV. Systematic method

In the following we describe a systematic method for the single trial estimation of multi channel evoked potentials measurements.

1. Measure the evoked potentials using multiple channels. Concatenate the channels as in (9). Form the matrix $z = (z_1, \ldots, z_N)$ of all measurements. Each column z_i is now concatenation of the evoked potential measurements in different channels corresponding the same stimuli. From the background EEG form a matrix $v = (v_1, \ldots, v_N)$ where the columns of the matrix are concatenations of the background measurements. The vectors v_i are typically measured before the stimulus. The time between the repetitions of the stimulus should be long enough and random to avoid the late potentials to corrupt the background estimate and locking of the background to the stimulus.
2. Calculate $C_v \approx N_v^{-1} \sum v_i v_i^T$. This is an estimate for the background covariance. Other methods are also applicable for the estimation of the covariance. The background can be modeled e.g. as an AR model and the covariance can then be calculated using the spectrum of the model. If you do not want to take the background EEG into account in the estimation simply replace $C_v \equiv I$ in the equations.
3. Form the matrix H with some generic way. A set of Gaussian shaped vectors with different delays and pre-selected shape is a suitable choice. If the measurements have trends, it is possible to include the constant and the first order polynomial basis vectors as columns of H.
4. Calculate $R_z \approx N_z^{-1} \sum z_i z_i^T$. This is an estimate for the correlation matrix of the measurements. Note that when the correlation matrix is used here, an approximation for the mean of the evoked potentials is modeled automatically as the first eigenvector of the correlation matrix.
5. Solve the ordinary eigendecomposition $R_z U = U\Lambda$ of the correlation matrix. The solution of only the principal eigenspace is also possible.
6. Form $H_{\mathcal{S}} = (u_1, \ldots, u_p)$ where u_i are column eigenvectors of R_z. The natural choice is to use the p eigenvectors that are associated with the p largest eigenvalues. The eigenvectors represent the common shape of evoked potentials in different channels and regularize the estimates to that direction. Due limited amount of data on which the determination of the eigenvectors is based, we know that any sharp spikes that occur in the vectors more likely have their origin in the noise than in true the evoked potentials. To minimize the disturbance caused by the sharp spikes the estimated eigenvectors can be smoothed e.g using the so-called smoothness priors approach [9]. The

regularization matrix can now be written in form

$$H_{\mathcal{S}} = (I + \gamma D_d^T D_d)^{-1}(u_1, \ldots, u_p) \qquad (14)$$

in which the regularization matrix D_d is usually second or third order difference matrix. A suitable value for the regularization parameter γ can often be selected by visual inspection of the smoothed eigenvectors.

7. Fix the regularization parameter α. This can be based on experience. The parameter α has to be tuned so that the estimate does not follow spurious peaks.
8. Calculate the estimate for the evoked potentials $\hat{s}$ with the equation

$$\hat{s} = H(H^T C_v^{-1} H + \alpha^2 H^T (I - H_{\mathcal{S}} H_{\mathcal{S}}^T) H)^{-1} H^T C_v^{-1} z. \qquad (15)$$

The selection of the basis vectors in H can be different, but the idea is that the basis is quite general, that is, several different types of possible measurements can be modeled with the selected basis. In subspace regularization s is regularized towards the null space of the regularization matrix $H_{\mathcal{S}}$. Our approach is to use all the available prior information in the construction of the matrix $H_{\mathcal{S}}$.

V. Case studies

In this section we compare the proposed multi channel method with the single channel method introduced in [4]. For reliable comparison we simulated data corresponding multi channel evoked potential measurements. Both of the methods were then applied to this simulated data and also to three sets of real measurements. Although the proposed methods are capable to estimate different kinds of evoked potentials we are mainly concentrated to the estimation of the P300 peak. The P300 peak is one of the most extensively studied cognitive potentials [13] and there exists many works where the trial-to-trial variation of the component is discussed, e.g [14], [15], [16].

A. *Simulation of the multi channel evoked potential measurements*

An effective way to simulate multi channel evoked potential measurements is to use a model for the electrical properties of the head and calculate potential distributions on the scalp caused by dipoles inside the head. This is known as forward problem in the context of source localization of EEG. This kind of simulation method has been discussed e.g. in [17], [18]. The real source of P300 peak is simultaneous activation of number of brain structures [19] and can not be fully described by some small set of dipoles. Nevertheless this method is a good way to produce simulated measurements which properties can be fully controlled and which are reasonably realistic especially in the sense of spatial correlation of the channels.

Our 3D finite element head model is based on segmentation of a MRI image of a human head. Two dipoles are placed inside the head model, see Fig. 4. Time variation of the dipoles, in fixed locations, is generated by multiplying magnitude of the dipoles by Gaussian shaped functions,

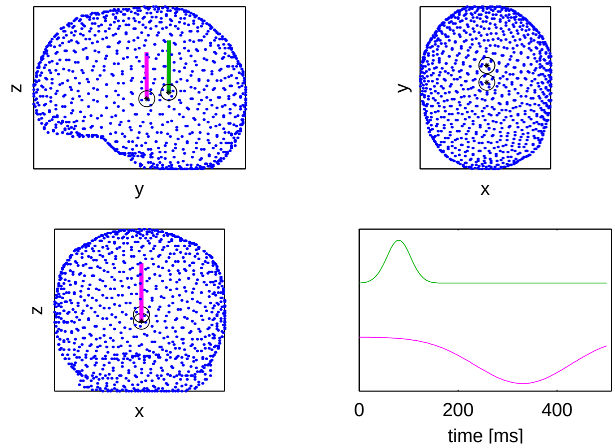

Fig. 4. Two dipoles used in the simulations. In lower right corner are magnitude coefficients as function of time.

functions are shown in lower right corner of Fig. 4. The potential distribution in the scalp is then calculated using the Finite Element Method in each time instant. Simulation of one ERP measurement necessitates then the solution of the scalp potential distribution 125 times if the simulated sampling rate is 250 Hz and the length of the ERP is 500 ms. The latency and the amplitude of the Gaussian humps varied between the stimuli for the simulation of the variation in the peaks of the ERPs. Variation was uniformly distributed and the amplitude and the latency of P300 peak were allowed to vary in larger scale than the ones in N1 peak.

Real EEG measurements were used as background noise in the simulations. The EEG measurements were from the test in which the P300 responses were recorded using the odd-ball paradigm and auditory stimuli. From this data we split out the pre-stimulus data of the standard stimuli, i.e. 500 - 0 ms before the stimuli. From this set of background EEG measurements we selected those which had no significant alpha rhythm or other disturbances. Then we summed this set of data to the simulated responses.

B. *Analysis of the simulations*

Single trial estimates for the simulated data were calculated using both of the proposed methods. In Fig. 5 single channel single trial estimates for four randomly selected stimuli in three channels are shown with noisy and noiseless simulations. We calculated the estimates with multi channel approach taking into account three channels (FZ, CZ, PZ). Estimates for the same stimuli as in Fig. 5 are show in Fig. 6. It can be seen that both of the methods give estimates that are reasonable good for each separate channel. The multi channel method gives somewhat smoother estimates especially for the channel PZ. The main difference in the performance of the methods is clearly seen in the N1 peak of the 3rd and the 36th stimulus. It can be seen that if the data has e.g. baseline disturbances the multichannel method gives estimates which are much closer to the true simulated potential. The single channel estimates tend to follow the biased data more closely. It is thus obvious that e.g. the amplitude esti-

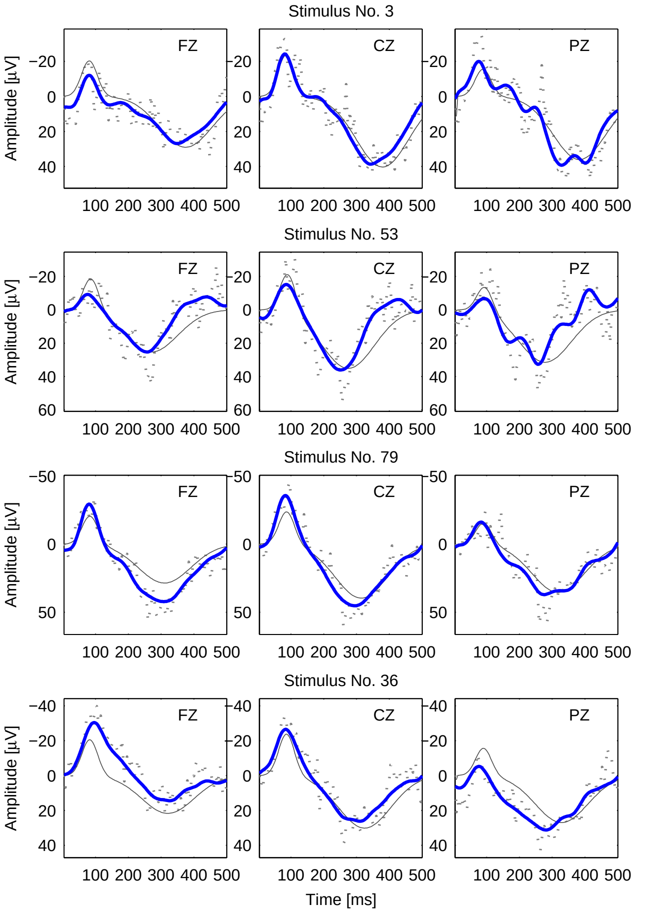

Fig. 5. Estimates (thick) for four randomly selected simulated stimuli in three channels with noisy (dotted) noiseless (thin) simulations. Estimates are obtained with single channel approach.

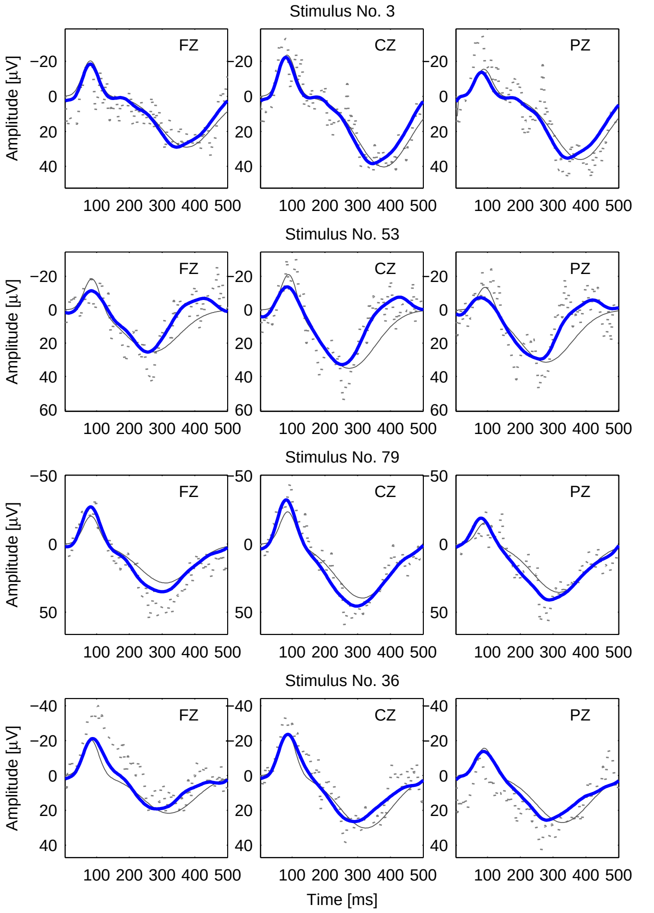

Fig. 6. Estimates (thick) for four randomly selected simulated stimuli in three channels with noisy (dotted) noiseless (thin) simulations. Estimates are obtained with multi channel approach.

mates that are obtained with the multi channel method are more comparable between the channels.

To demonstrate this phenomenon we also estimated the latencies and amplitudes of the P300 peaks. The peak estimation was based on the fitting of the second-order polynomial to the estimate in the vicinity of the peak. The method is introduced in [20]. Mean and standard deviation of the simulated responses calculated from peak estimates are shown in Table I with the corresponding values calculated from the estimated potentials. The latencies and the amplitudes estimated from the averages are also shown in the table. It can be seen that both methods can estimate the latency equally well and also the amplitude estimates are equivalent.

The main difference of the methods is in ability to preserve amplitude ratios between different channels. This can be seen from Table II where mean and standard deviation of amplitude differences of between CZ and FZ and between CZ and PZ are shown. In the table percentage of the potentials for which the difference between the amplitudes of the peaks on different channels is greater than zero are shown. The multi channel approach clearly preserves amplitude ratios better.

TABLE I

MEAN ($\overline{x} \pm \sigma_{\overline{x}}$) AND STANDARD DEVIATION (S.D.) FOR LATENCY AND AMPLITUDE OF BOTH SIMULATED (SIM.) AND ESTIMATED (SC, MC, AVG.) P300 PEAKS FOR THREE CHANNELS. SC REFERS TO SINGLE CHANNEL METHOD AND MC TO MULTI CHANNEL METHOD AND AVG. TO TRADITIONAL AVERAGE.

		Latency [ms]			Amplitude [μV]		
Ch.		$\overline{x}$	$\sigma_{\overline{x}}$	S.D.	$\overline{x}$	$\sigma_{\overline{x}}$	S.D.
FZ	Sim.	326	4	39.6	24.9	0.3	2.9
	sc	320	3	33.6	26.0	0.6	6.3
	mc	318	3	33.0	25.2	0.5	5.3
	Avg.	324	-	-	24.1	-	-
CZ	Sim.	330	4	39.6	34.6	0.4	4.1
	sc	318	3	35.3	34.5	0.7	7.0
	mc	318	3	33.8	33.1	0.6	6.6
	Avg.	324	-	-	31.9	-	-
PZ	Sim.	334	4	39.6	31.0	0.3	3.6
	sc	318	3	35.9	30.4	0.7	7.0
	mc	317	3	34.7	29.7	0.5	5.8
	Avg.	324	-	-	28.5	-	-

TABLE II
MEAN ($\overline{x} \pm \sigma_{\overline{x}}$) AND STANDARD DEVIATION (S.D.) OF AMPLITUDE DIFFERENCES [μV] OF P300 PEAK IN SIMULATED DATA. SIM. REFERS TO SIMULATED RESPONSES, SC TO THE ESTIMATED POTENTIALS OBTAINED WITH SINGLE CHANNEL APPROACH AND MC TO ESTIMATES OBTAINED WITH MULTI CHANNEL APPROACH. COLUMN > 0 SHOW THE PERCENTAGE OF DIFFERENCES WHICH ARE GREATER THAN ZERO.

		$\overline{x}$	$\sigma_{\overline{x}}$	S.D.	> 0 [%]
CZ-FZ	Sim.	9.69	0.11	1.14	100
	sc	8.5	0.4	4.53	96.36
	mc	7.92	0.15	1.59	100
CZ-PZ	Sim.	3.63	0.04	0.43	100
	sc	4.1	0.5	4.82	80.91
	mc	3.38	0.13	1.39	99.1

C. Real data

To evaluate the performance of the methods for the real measurements the P300 responses of three test persons (A,B,C) were recorded using the odd-ball paradigm and auditory stimuli. We calculated estimates for the evoked potentials with both methods. In the multi channel estimation we used three channels (FZ,CZ,PZ). Estimates for four randomly selected stimuli in three channels for person A are shown in Fig. 7 for the single channel method and in Fig. 8 for the multi channel method. The effect of the multi channel approach in the estimates can be seen e.g. in estimates for stimulus 28. In the multi channel estimate the amplitude of the P300 peak is smaller in FZ and larger in PZ. Same kind of behavior can also be seen in the estimates for stimulus 19. In the estimates for stimulus 36 the multi channel method does not follow the disturbance peak in channel FZ as much as single channel method.

The latencies and amplitudes of P300 peaks were also estimated. Mean and standard deviation of latency and amplitude estimates of P300 peak are show in Table III for test persons A-C. The latency and amplitude estimates are calculated also using average of measurements. It can be seen that there are no big differences in the means and the standard deviations of the latencies and the amplitudes between the methods. Also the amplitude differences between the channels are comparable to that calculated using the average of the measurements. The main difference of the methods is in the estimation of the amplitude differences between the single trials. The histograms of the amplitude differences between the channels CZ and FZ as well as between the channels CZ and PZ are presented in the Fig. 9 for the single channel method and in the Fig. 10 for the multi channel method. There is much less variation in the amplitude differences calculated using the multi channel method. This is closer to the true situation at least when the topology of the neural source of the potential is rather compact and if the variation of the amplitudes between the trials is mainly due the variation of the strength of the source. The same effect can also be seen in the Table IV for the test persons A-C.

TABLE III
MEAN ($\overline{x} \pm \sigma_{\overline{x}}$) AND STANDARD DEVIATION (S.D.) FOR LATENCY AND AMPLITUDE OF P300 PEAKS OF ESTIMATED POTENTIALS FOR TEST PERSONS A-C. LATENCY AND AMPLITUDE ESTIMATES CALCULATED FROM THE AVERAGE (AVG.). SC REFERS TO SINGLE CHANNEL METHOD AND MC TO MULTI CHANNEL METHOD.

			Latency [ms]			Amplitude [μV]		
	Ch.		$\overline{x}$	$\sigma_{\overline{x}}$	S.D.	$\overline{x}$	$\sigma_{\overline{x}}$	S.D.
Test Person A	FZ	sc	316	4	32.5	21.7	1.8	13.2
		mc	317	4	28.4	21.2	1.6	12.1
		Avg.	316	-	-	20.1	-	-
	CZ	sc	319	5	40.1	26.0	2.0	15.0
		mc	319	4	31.8	25.3	1.9	13.8
		Avg.	316	-	-	23.6	-	-
	PZ	sc	338	6	45.8	24.2	1.7	12.4
		mc	330	5	35.6	22.9	1.6	12.0
		Avg.	328	-	-	21.6	-	-
Test Person B	FZ	sc	315	5	35.6	19.8	1.3	9.33
		mc	320	4	31.0	18.4	1.1	8.07
		Avg.	312	-	-	18.5	-	-
	CZ	sc	322	5	36.9	23.5	1.4	10.4
		mc	322	5	34.7	21.9	1.3	9.51
		Avg.	316	-	-	21.7	-	-
	PZ	sc	336	5	38.4	18.3	1.3	9.47
		mc	336	5	36.2	17.6	1.1	7.93
		Avg.	332	-	-	17.3	-	-
Test Person C	FZ	sc	327	3	22.7	16.7	0.9	7.4
		mc	335	3	28.0	16.2	0.8	6.6
		Avg.	332	-	-	17.3	-	-
	CZ	sc	340	5	40.0	24.6	1.2	10.5
		mc	336	3	23.4	23.1	1.1	9.78
		Avg.	336	-	-	24.2	-	-
	PZ	sc	354	4	29.9	27.6	1.4	11.7
		mc	351	3	24.4	25.5	1.3	10.7
		Avg.	356	-	-	26.3	-	-

TABLE IV
MEAN ($\overline{x} \pm \sigma_{\overline{x}}$) AND STANDARD DEVIATION (S.D.) OF AMPLITUDE DIFFERENCES OF P300 PEAK FOR TEST PERSONS A-C. SC REFERS TO SINGLE CHANNEL METHOD AND MC TO MULTI CHANNEL METHOD.

			$\overline{x}$	$\sigma_{\overline{x}}$	S.D.	> 0 [%]
A	CZ-FZ	sc	4.3	0.7	5.3	80.0
		mc	4.1	0.3	2.1	100
	CZ-PZ	sc	1.7	0.8	5.6	61.8
		mc	2.4	0.3	2.0	81.8
B	CZ-FZ	sc	3.7	0.7	5.6	78.2
		mc	3.5	0.3	2.0	96.4
	CZ-FZ	sc	5.2	0.9	6.6	76.4
		mc	4.3	0.4	3.0	96.4
C	CZ-FZ	sc	7.8	0.7	5.9	93.2
		mc	6.8	0.4	3.5	98.7
	CZ-PZ	sc	-3.0	0.6	5.0	31.1
		mc	-2.45	0.13	1.1	1.4

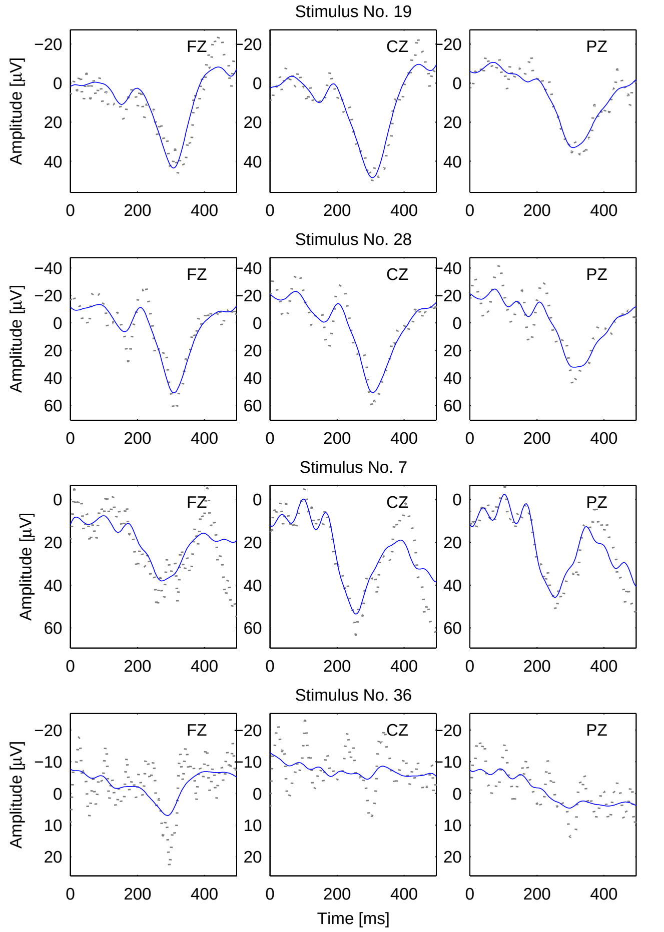

Fig. 7. Single channel single trial estimates (thin) for four randomly selected stimuli in three channels for test person A. The measurements are shown as dotted lines.

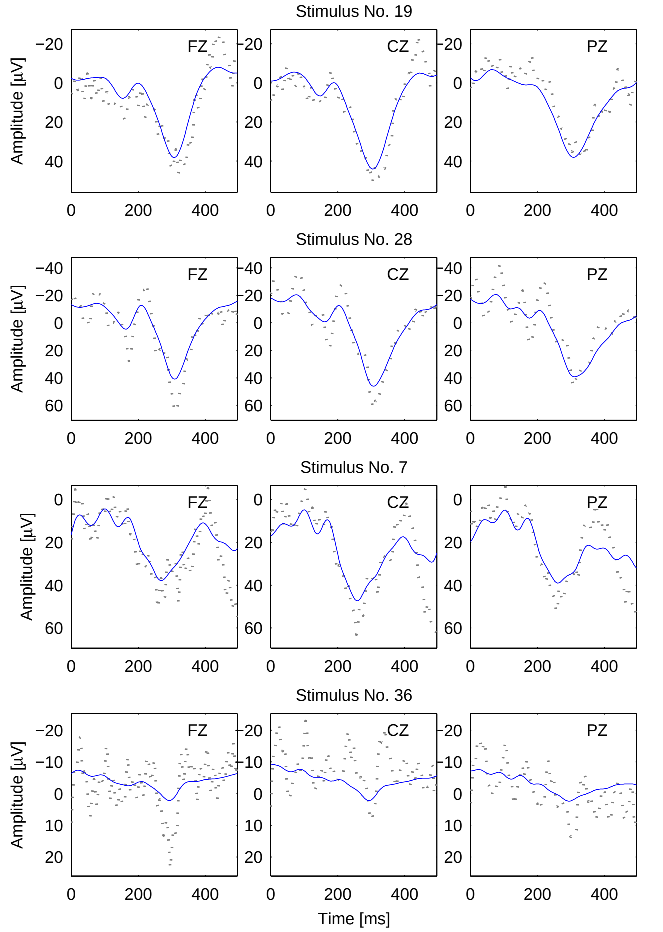

Fig. 8. Multi channel single trial estimates (thin) for four randomly selected stimuli in three channels for test person A. The measurements are shown as dotted lines.

VI. Discussion

We have extended the single channel single trial estimation method introduced in [4] for multi channel measurements. In the multi-channel method the spatial correlation of the measurements is taken into account by using the subspace regularization method. This approach does not necessitate the proper modeling of the head as volume conductor which keeps the proposed method easy to implement.The main benefit of the proposed method is it's ability to preserve the amplitude ratios of the channels in the estimates.

The most promising application of the single trial estimation methods is the ability to study e.g. habituation of the potentials during long tests. Usually habituation affects especially to the amplitude of the potentials. Rating of the psychological importance of the stimulus will also necessitate exact amplitude information about the potentials. Hence the method that is proposed in this article has many applications. The preservation of the amplitudes is also important in the future if one aims to use the single trial estimates in the source localization or cortical imaging applications [21]. The use of single trials in topographic estimates is suggested also in [22].

References

[1] E. John, D. Ruchkin, and J. Villegas, "Experimental background: signal analysis and behavioral correlates of evoked potential configurations in cats," *Ann N Y Acad Sci*, vol. 112, pp. 362–420, 1964.

[2] K. Yu and C. McGillem, "Optimum filters for estimating evoked potential waveforms," *IEEE Trans Biomed Eng*, vol. 30, pp. 730–737, 1983.

[3] C. W. Groetsch, *Inverse Problems in the Mathematical Sciences.* Vieweg, 1993.

[4] P. Karjalainen, J. Kaipio, A. Koistinen, and M. Vauhkonen, "Subspace regularization method for the single trial estimation of evoked potentials," *IEEE Trans Biomed Eng*, vol. 46, pp. 849–860, July 1999.

[5] J. Westerkamp and J. Aunon, "Optimum multielectrode a posteriori estimates of single-response evoked potentials," *IEEE Trans Biomed Eng*, vol. 34, pp. 13–22, 1987.

[6] P. A. Karjalainen, J. P. Kaipio, V. Kolehmainen, A. S. Koistinen, and J. Partanen, "Regularization approach to single trial estimation of multi channel evoked potentials," in *Proceedings of the First Joint BMES/EMBS Conf*, vol. 1, (Atlanta), p. 423, October 1999.

[7] P. Brown, *Measurement, Regression, and Calibration.* Oxford Science Publications, 1993.

[8] M. Hanke and P. Hansen, "Regularization methods for large-scale problems," *Surv Math Ind*, vol. 3, pp. 253–315, 1993.

[9] W. Gersch, "Smoothness priors," in *New Directions in Time Series Analysis, Part II*, pp. 113–146, Springer-Verlag, 1991.

[10] A. Tarantola, *Inverse Problem Theory.* Elsevier, 1987.

[11] P. Karjalainen, *Regularization and Bayesian methods for evoked potential esimation.* PhD thesis, University of Kuopio, Department of Applied Physics, 1997.

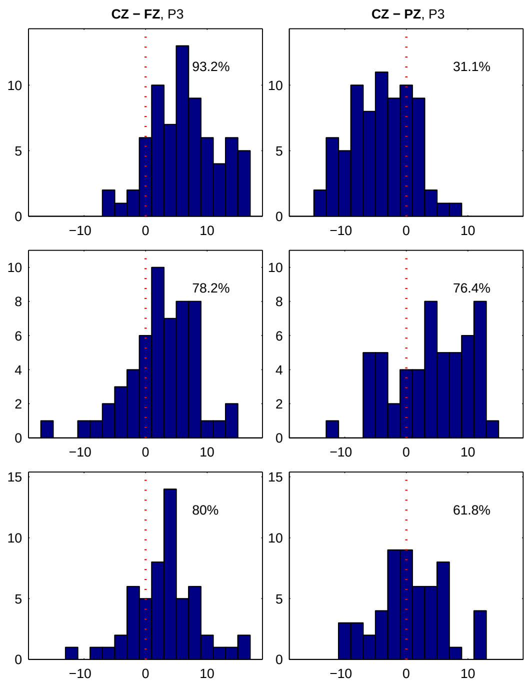

Fig. 9. Histograms of amplitude differences of P300 peaks in peak estimates calculated from single channel evoked potential estimates. In the first row for test person A, second for B and third for C. Percent number in the figures shows percentage of differences which are greater than zero.

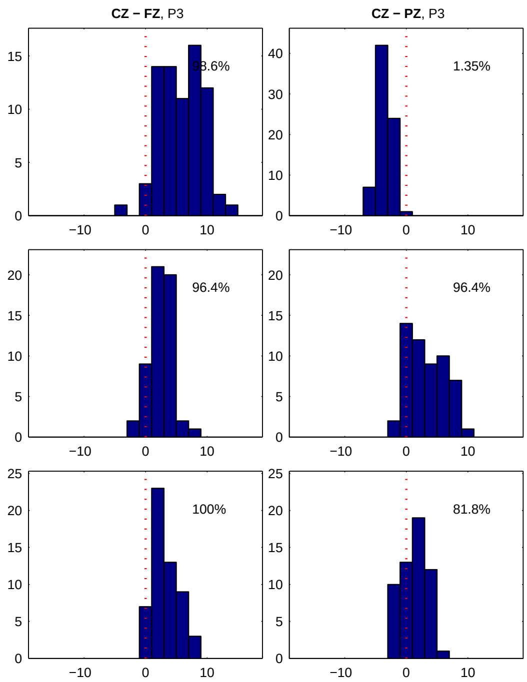

Fig. 10. Histograms of amplitude differences of P300 peaks in peak estimates calculated from multi channel evoked potential estimates. In the first row for test person A, second for B and third for C. Percent number in the figures shows percentage of differences which are greater than zero.

[12] G. Wahba, *Spline Models for Observational Data.* SIAM, 1990.
[13] K. Pilgreen, "Physiologic, medical and cognitive correlates of electroenceohalography," in *Neocortical Dynamics and Human EEG Rhytms* (P. Nunez, ed.), ch. 4, pp. 195–248, Oxford University Press, 1995.
[14] M. Brazier, "Evoked responses recorded from the depths of the human brain," *Ann N Y Acad Sci*, vol. 112, pp. 33–59, 1964.
[15] W. Lammers and P. Badia, "Habituation of P300 to target stimuli," *Physiol Behav*, vol. 45, pp. 595–601, 1989.
[16] A. Gevins and B. Cutillo, "Neuroelectric measures of mind," in *Neocortical Dynamics and Human EEG Rhytms* (P. Nunez, ed.), ch. 7, pp. 304–338, Oxford University Press, 1995.
[17] Y. Yan, P. Nunez, and R. Hart, "Finite-element model of the human head: scalp potentials due to dipole sources," *Med Biol Eng Comput*, vol. 29, pp. 475–481, 1991.
[18] J. Le and A. Gevins, "Method to reduce blur distortion from EEG's using a realistic head model," *IEEE Trans Biomed Eng*, vol. 40, pp. 517–528, 1993.
[19] P. Nunez, *Electric Fields of the Brain: The Neurophysics of EEG.* Oxford University Press, New York, 1981.
[20] C. McGillem, K. Yu, and D. Childers, "Effects of ongoing EEG on latency measurements of evoked potentals," in *Proc. 4th Annual Conf. IEEE-EMBS*, (Philadelphia), pp. 60–63, 1982.
[21] R. Sencaj and J. Aunon, "Dipole localization of average and single visual evoked potentials," *IEEE Trans Biomed Eng*, vol. 29, no. 1, pp. 26–33, 1982.
[22] D. Liberati, L. Narici, A. Santonini, and S. Cerutti, "The dynamic behavior of the evoked magnetoencephalogram detected through parametric identification," *J Biomed Eng*, vol. 14, pp. 57–63, 1992.