Comparison of amplitude estimates in the single trial estimation of evoked potentials

P.O. Ranta-aho[1], P.A. Karjalainen[1], A.S. Koistinen[2], J.Kaipio[1], J.Partanen[3]
[1] University of Kuopio, Department of Applied Physics, Kuopio, Finland
[2] Finnish Institute of Occupational Health, BrainWork Laboratory, Helsinki, Finland
[3] Kuopio University Hospital, Department of Clinical Neurophysiology, Kuopio, Finland

Abstract—**A method for the single trial estimation of the evoked potentials is presented. The method is based on the regularized least squares scheme. An approach to single trial estimation of multi channel evoked potentials is derived by applying additional information of the spatial correlation of the channels to the estimation. Amplitude estimates obtained with these methods are compared. The performance of the methods is evaluated using real data of target responses in the P300 test. 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, regularization**

I. INTRODUCTION

The goal in the evoked potential analysis is currently often the estimation of single potentials, a task, what we call single trial estimation. Common approach to achieve this goal 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 the very low signal-to-noise ratio.

One possibility to improve the estimates is to apply additional information about the potentials to the estimation. Evoked potential can be assumed to be smooth or we may assume some limits for the possible locations of the peaks in the potentials. A time varying filter introduced in [1] is of this of estimator. An interesting way to add temporal information to estimation is to use the so-called regularized least squares method. The method has been used in [2] for case of single channel measurements. In the case of multi channel measurements the spatial correlation between the channels can also be interpreted in the form of additional information. A method for this task is introduced in [3]. A regularized least squares approach has been used in [4]. Many benefits of the multi channel approach is discussed in [5].

In this paper we have compared the single channel and the multi channel approaches. Especially the amplitude estimates in the P300 measurements are studied. Particularly the capability to preserve the order of the amplitudes of P300 peaks between different channels has been of the special interest.

II. LINEAR OBSERVATION MODEL

We denote the sampled potential measurement after i'th stimulus with a column vector z_i. Conventionally the so-called additive noise model is used for the measurements

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

The electrical activity of the brain is the source for both s_i and e_i. The evoked potential s_i corresponds to the part of the activity that is correlated with the stimulus. It is thus 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.

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 θ_i is a length j vector of parameters and H is a matrix that contain the basis vectors ψ_j in its columns. The observation model is now

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

A. Selection of the basis

With linear observation model the choice of the observation matrix H has a significant role. Obviously the best choice would be the true physical model. This would however require the modeling of electrical properties of the head, which is not a trivial task. We will use a simplier approach here.

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. It would also be 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.

Both the use of generic basis vectors and the principal component regression approach have their own benefits. Generic vectors are usually robust and easy to generate. They also usually model different intervals of the measurement vector in a homogeneous way. The eigenvector basis is on the other hand 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 based on the measurements z_i. We state a generalized least squares solution for the parameters θ_i

$$\hat{\theta}_i = \arg\min_{\theta_i} \left\{ \|(z_i - H\theta_i)\|^2 + \alpha^2 \|L\theta_i\|^2 \right\}. \tag{4}$$

The statement $\arg$ means the argument of the expression. The solution (4) is called the generalized Tikhonov regularized solution and it is clearly a modification of the ordinary weighted least squares solution $\hat{\theta}_{\text{LS}} = \arg\min_{\theta_i} \left\{ \|(z_i - H\theta_i)\|^2 \right\}$ to the direction in which the norm $\|L\theta_i\|$, the so-called *side constraint*, gets smaller.

The regularization matrix L can be e.g. the second derivative approximation. 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 that the regularized solution can be written in the form

$$\hat{\theta}_i = (H^T H + \alpha^2 L^T L)^{-1} H^T z_i \tag{5}$$

III. Single channel approach

In this section we show how the principal component approach can be combined with the generic basis vectors using the regularized least squares method. First we use a generic Gaussian basis in the columns of H. Next we calculate the eigenvectors of the correlation matrix R_z and use p first eigenvectors as columns of a matrix $H_{\mathcal{S}}$. $H_{\mathcal{S}}$ will now contain an orthonormal basis of the subspace $\mathcal{S}$. Now we want that the estimated evoked potential is close to this subspace. The projection of $s_i = H\theta_i$ onto $\mathcal{S}$ is $(H_{\mathcal{S}} H_{\mathcal{S}}^T)H\theta_i$. The distance of s_i from $\mathcal{S}$ can then 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 $\|L\theta_i\|$ is small for all expectable θ_i, we thus select $L = (I - H_{\mathcal{S}} H_{\mathcal{S}}^T)H$. Since

$$\begin{aligned} L^T L &= H^T (I - H_{\mathcal{S}} H_{\mathcal{S}}^T)^T (I - H_{\mathcal{S}} H_{\mathcal{S}}^T) H \\ &= H^T (I - H_{\mathcal{S}} H_{\mathcal{S}}^T) H, \end{aligned} \tag{6}$$

the desired solution for the parameters θ_i can be written in the form

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

The estimate for the evoked potential is then

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

This is called the subspace regularized solution[6].

So 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. Set of eigenvectors of the covariance matrix of the measurements is used to represent the correlation between the set of measurements or in the multi channel case the spatial information about the problem. These are the columns of the matrix $H_{\mathcal{S}}$. The parameter α controls the weight between the different sets.

IV. Multi channel approach

In the chase of multi channel measurements we use the following notation

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

That is, z_i is a concatenate of the measurements of M channels.

Although the evoked potential and the background EEG are independent both the activities s_i and e_i are highly correlated between different channels $j = 1, \ldots, M$. In the case of multi channel measurements we can combine the spatial information between the channels by using the subspace regularization method.

We use first a generic Gaussian basis for each channel. That is, the columns of H_j are Gaussian shaped vectors. The channels are modeled separately, that is, the matrix H is a block diagonal matrix.

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

This model does not contain any dependence between the channels. We form next 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 first few eigenvectors to form the regularization matrix $H_{\mathcal{S}}$. Because the correlation is calculated using the stacked measurements z, the eigenvectors model also the correlation between the separate channels. The estimate is again obtained with 7 and 8.

V. Choice of the regularization parameter

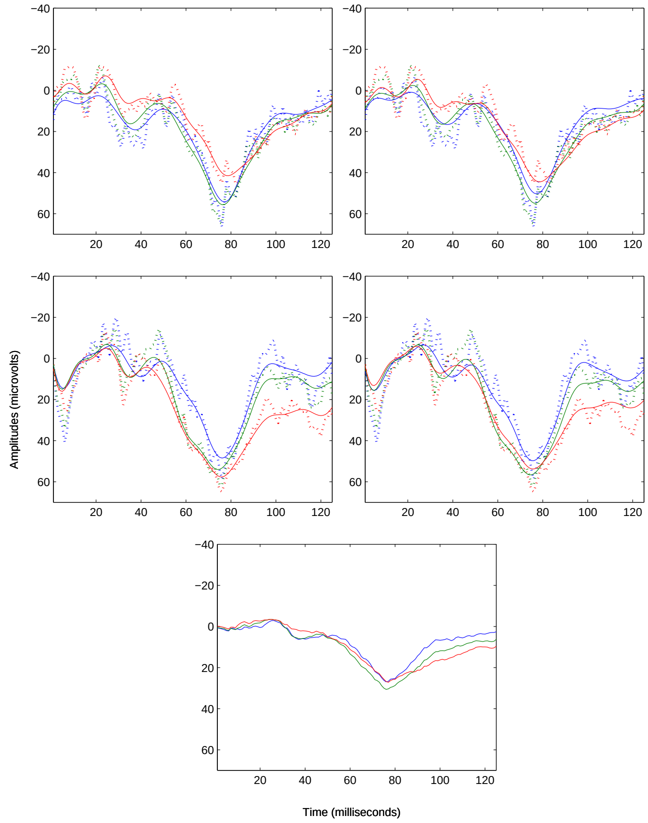
There exist several methods for estimating the optimal value of the regularization parameter α. The so-called GCV method [7] is one of the possibilities. Using GCV in selection of the regularization parameter for single channel approach is discussed in [2]. However, the proposed method is quite robust to small changes of α. So here our selection of regularization parameter is based on experience, value $\alpha = 1.5$ is fixed for both the single and multi channel approach.

Typical values of α are from 0.5 to 5. If selected value is too small the estimation will be non smooth and follow data closely. Too large value will force the estimation be near constant zero.

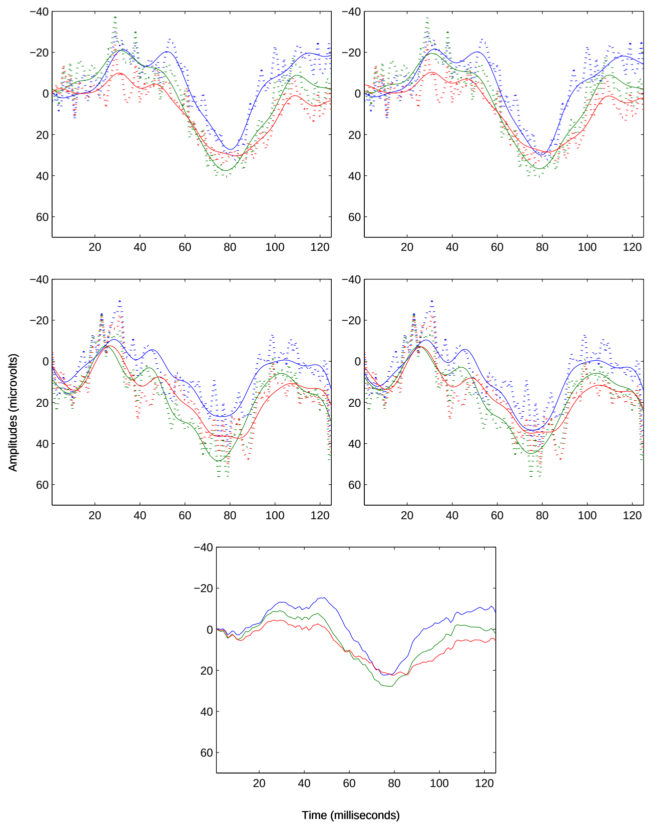
VI. Comparison of the two methods

The P300 responses of two test persons were recorded using the odd-ball paradigm and auditory stimuli. The single channel and the multi channel approaches were then applied to the target stimuli. Three channels (FZ,CZ,PZ) were used.

The brain activity measured as P300 peak can be assumed to appear in the same location in the brain for all stimuli. Theoretically the relations between the measured amplitudes of the peaks on different channels should then be constants as the function of stimulus. However the trends and disturbances caused by the background EEG can differ between the channels. The contribution of these disturbances is not easy to be taking into account in the single channel approach. In the multi channel approach the estimates between the channels are binded together by the side constraint term. This approach should then better preserve the amplitude information between



(a)



(b)

Fig. 1. Two selected estimated (solid) and measured (dotted) target responses for two test persons (a) and (b). The estimates in the left columns are calculated with single channel approach, in the right ones with multi channel approach. In the last row is the average of all measurements. Blue represents channel FZ, green CZ and red PZ.

the measured channels. The measurements and the estimates are shown for two selected stimuli in Fig 1.

We calculated also the differences in the amplitudes between the channels FZ and CZ, and between CZ and PZ for all target responses. The estimated differences are shown in Fig. 2 with the corresponding mean and standard deviation. The corresponding histograms are shown in Fig. 3. It can be seen that the standard deviation of the differences is smaller in the multi channel approach. This is true also for the standard deviation of the estimated latencies of the peaks although the difference between methods seems to be much smaller. However the mean of the differences is not necessarily closer the difference of amplitudes measured from the averages. From the histograms shown in Fig 3 it can be seen that the multi channel approach preserves the order of the peak amplitudes between the channels better.

VII. Conclusions

The multi channel approach for the single trial estimation is useful if the amplitude information is of the interest. The capability of the estimation method to preserve the amplitude information is extremely important in the future if one is interested in to use the estimates in the source localization or in cortical imaging applications.

References

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

[2] 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.

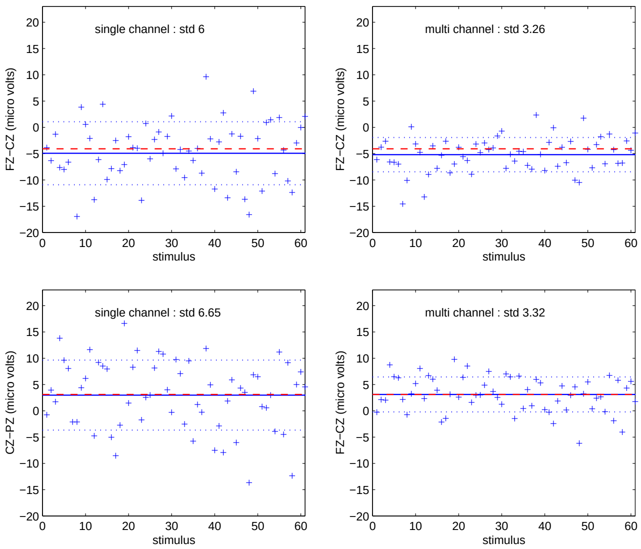
[3] J. Westerkamp and J. Aunon, “Optimum multielectrode a posteriori estimates of single-response evoked potentials,” *IEEE Trans Biomed Eng*, vol. 34, pp. 13–22, 1987.

[4] 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.

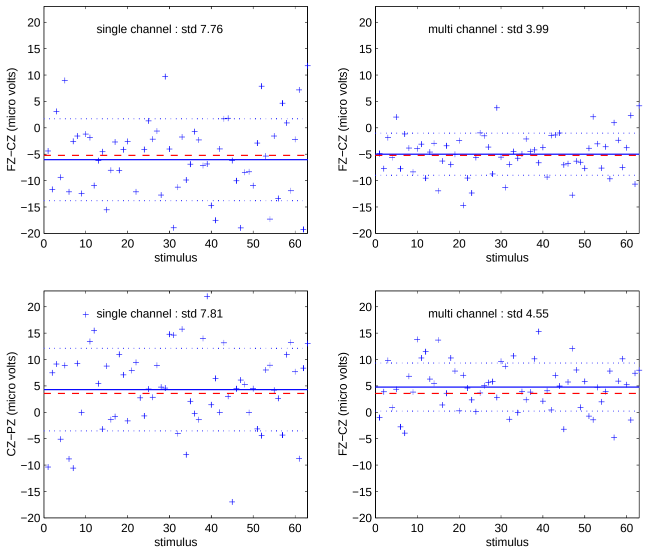
[5] A. Geva, H. Pratt, and Y. Zeevi, “Multichannel wavelet-type decomposition of evoked potentials: model-based recognition of generator activity,” *Med Biol Eng Comput*, vol. 35, pp. 40–46, 1997.

[6] M. Vauhkonen, D. Vadász, P. Karjalainen, E. Somersalo, and J. Kaipio, “Tikhonov regularization and prior information in electrical impedance tomography,” *IEEE Trans Med Imaging*, vol. 17, pp. 285–293, 1998.

[7] G. Wahba, *Spline Models for Observational Data*. SIAM, 1990.

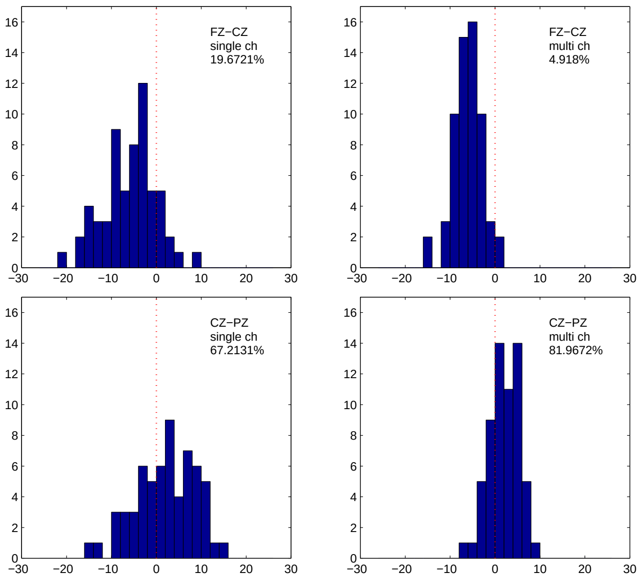


(a)

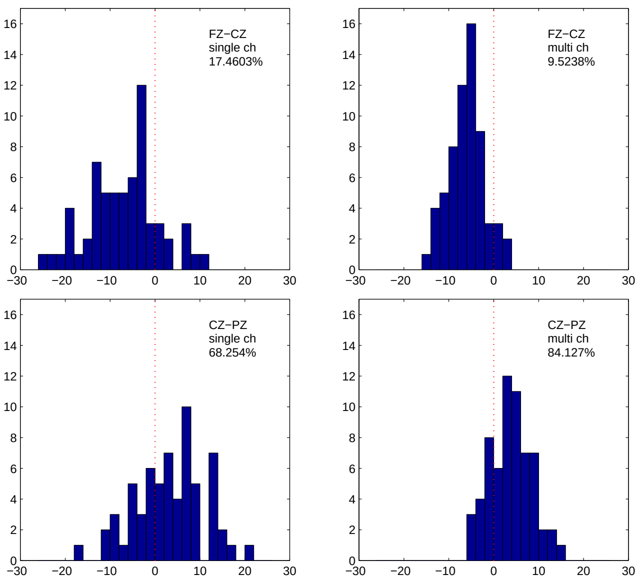


(b)

Fig. 2. In the first row is the difference of amplitudes between channels FZ and CZ and in the second row between CZ and PZ. The left columns are for the single channel estimation and the right ones for multi channel estimation. The solid blue line is average of differences, the blue dashes show standard deviation. The red dash is difference of amplitudes calculated from average of measurements.



(a)



(b)

Fig. 3. The histograms of differences shown in Fig 2. In the left columns for the single channel estimation and in the right ones for multi channel estimation. The percent values represents the positive part of the histogram. The red dash shows the zero.