University of Kuopio Department of Applied Physics Report Series ISSN 0788-4672

Estimation of the Dynamics of Event-related Desynchronization

J.K. Hiltunen, P.A. Karjalainen, J. Partanen,
J.P. Kaipio

February 26, 1998

Report No. 3/98

This manuscript has been submitted to *Med. Biol. Eng. Comput.*

University of Kuopio • Department of Applied Physics
P.O.Box 1627, FIN-70211 Kuopio, Finland

Estimation of the Dynamics of Event-related Desynchronization*

J.K. Hiltunen, P.A. Karjalainen, J. Partanen, J.P. Kaipio

February 26, 1998

Abstract A method for the estimation of medium rate transitions of nonstationary EEG is proposed. The method is applicable to such EEG dynamics that are between a) fast transitions for which segmentation procedures are used and b) slow transitions for which adaptive filters work properly. The estimation of the transition dynamics is based on a novel time-varying autoregressive model. The performance of the method is evaluated with realistic simulations of known transition dynamics and it is shown to be able to track even slight medium-rate transitions. The method is then applied to the estimation of the dynamics of event–related desynchronization (ERD).

1 Introduction

The traditional approach in the analysis of background EEG has been the segmentation of EEG into (almost) stationary epochs [1, 2, 3]. This approach works often reasonably well in both very fast and very slow transition cases. In the former case a segmenter [4, 5, 6] and in the latter case an adaptive algorithm [7, 8] can be used. However, there are only few algorithms for the estimation of such transitions whose rate of change falls between these two limiting cases.

Recently there has been increasing interest in the transition dynamics of EEG, that is, how the change occurs – not only the initial and final steady states. The so-called ERD/ERS test (for definition, see below) is an example of an application in which the transition has been investigated for clinical relevance [9, 10, 11, 12] and methods for this task have been developed, e.g. in [13]. The ERD/ERS can be briefly described as follows.

Stimulus–induced blocking or attenuation of rhythms within the alpha frequency band (8–12 Hz) is called event–related desynchronization (ERD). The opposite phenomenon, event–related synchronization (ERS), is the increase of rhythms within the alpha band [14, 15]. Both ERD and ERS have been widely used in the assessment of lesions and neurophysiological pathologies such as strokes and tumors [16, 17]. However, also in these cases mainly the steady states before and after the transitions have been examined for the clinical evaluation. It is also believed that there might be different delays in the occurrence of the desynchronization after the stimulus, and that these delays might convey some information on the neurophysiological state.

The visual appearance and statistical properties of EEG in the synchronized and desynchronized states can be very different with different persons. The transition can be very fast and the statistical properties as well as visual appearance can differ considerably before and after the transition. In some cases of event–related desynchronization, a segmenter could be used for transition detection. However, event–related desynchronization is a narrow band process usually with considerable amount of energy in multiples of the main alpha frequency. In these cases the synchronized state exhibits amplitude modulation type characteristics, see for example Figs. 9–11. The co-occurrence of a suppression phase in the modulation and an ERD transition makes it difficult to estimate any parameters of the transition.

*J.K. Hiltunen, P.A. Karjalainenand J.P. Kaipio are with the Dept. of Applied Physics, University of Kuopio, P.O.Box 1627, FIN-70211, Kuopio, FINLAND. J.K. Hiltunen and J. Partanen are with the the Dept. of Clinical Neurophysiology, Kuopio University Hospital, P.O.Box 1777, FIN-70211, Kuopio, FINLAND. Correspondence to: J.P. Kaipio, Department of Applied Physics, University of Kuopio, P.O.Box 1627, 70211 Kuopio, FINLAND, email: kaipio@venda.uku.fi.

In other cases, the changes in the statistics of EEG may be too small to be detected properly with conventional segmenters. These problematic cases are discussed more in Section 3.2 in which the proposed method is applied to event–related desynchronization data.

The statistical properties of most previously used estimation methods such as [15] do not enable the estimation of a single sample. However, if for example trends in the transition delays in consecutive tests are to be detected we have to be able to estimate the individual responses with adequate accuracy.

An algorithm for the estimation of such transition dynamics in EEG was proposed in [13, 18]. This algorithm is based on the basis constrained least squares time–varying autoregressive model (TVARLS) and a principal component type approach in the construction of the basis functions. The model works relatively well in the estimation of the transition dynamics and it can take into account prior distribution of the initial and final states but it has the drawback that it has to initialized for each patient separately. This initialization needs supervision and is computationally burdensome. This is an unwanted requirement and an automatic algorithm that does not need supervision would be very desirable.

Due to the variations in the characteristics of both ERD and ERS states, this means that it is then practically impossible to implement statistical description of the ERD and ERS states into the algorithm.

In this paper we propose an algorithm for the estimation of medium rate EEG transition dynamics. This algorithm is based on a modification of the general TVARLS algorithm. This modification reparametrizes the TVARLS problem so that the transition time depends only on a single parameter of the model [19]. The estimation properties of the method are evaluated using realistic EEG simulations with known transition dynamics. We show that the proposed method conforms to the a priori knowledge of the process and is able to estimate a single transition within an approximately known region. Thereafter, the method is applied to such ERD data whose transitions are nontrivial.

The paper is organized as follows. In Section 2.1 the basis constrained time–varying autoregressive model is formulated and the basis selection problem is discussed in Section 2.2. The modification of the TVARLS problem is described in Section 2.3. The TVAR model is applied to simulated ERD in Section 3.1 and to ERD data in Section 3.2. Finally, in Section 4 we discuss some extentions to the proposed method.

2 Methods

2.1 The general TVARLS model

The time-varying AR(p) model (TVAR) is generally of the form

$$x_t = \sum_{k=1}^{p} a_k(t) x_{t-k} + e_t, \tag{1}$$

where p is the order of the model, e_t are the residuals and $a_k(t)$ are the time–varying prediction coefficients. The estimation of the coefficient evolutions $a_k(t)$ in (1) is a highly underdetermined problem and no meaningful estimates can be obtained without any further constraints or the implementation of a priori knowledge.

In the basis constrained TVAR problem the coefficients are constrained to

$$a_k(t) = \sum_{\ell=0}^{M} c_{k\ell} \phi_\ell(t) \ , \tag{2}$$

where $\phi_\ell(t),\ \ell = 0, \ldots, M$ are the basis functions.

The minimization of the 2-norm of the residuals in (1) with the constraints (2) leads to a quadratic problem, the TVARLS problem, with the $(M+1)p$ parameters $c_{k\ell}$. The TVARLS problem was introduced first in [20] and has thereafter been partially reformulated and applied

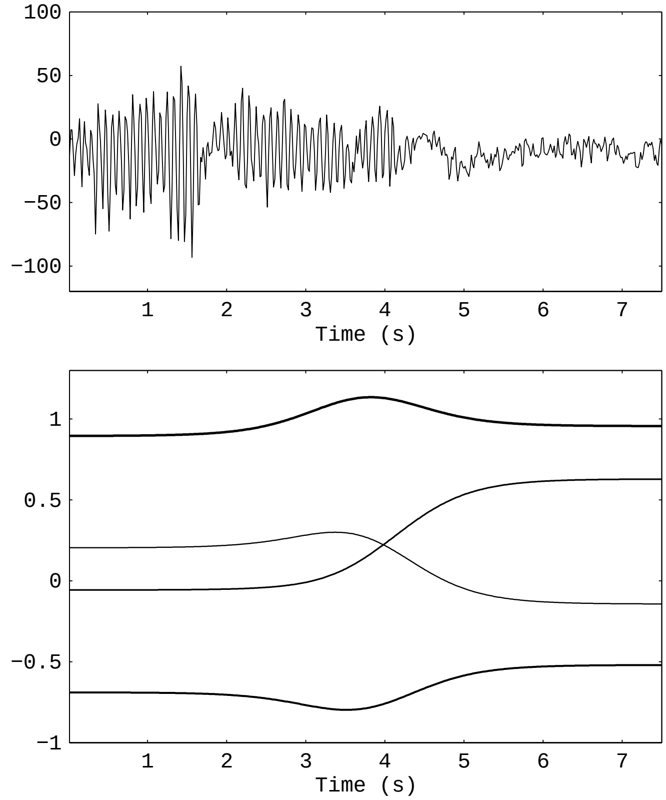

Figure 1: An example of ERD (top) and the estimated TVAR(4) time-varying coefficients $a_k(t)$, $k = 1, \ldots, 4$ (bottom).

to EEG and speech modelling for example in [21, 22, 23, 24, 25, 26, 27]. With the exceptions of [18, 19] the work in TVARLS modelling has been directed almost exclusively to the selection of the basis functions.

The traditional method for the solution of the constrained LS problem (1–2) is the covariance formulation [21]. The LS solution can be accessed more conveniently by writing $C = (c_{10}, \ldots, c_{1M}, \ldots, c_{p0}, \ldots, c_{pM})^{\mathrm{T}}$, $X = (x_{p+1}, \ldots, x_T)^{\mathrm{T}}$, $E = (e_{p+1}, \ldots, e_T)^{\mathrm{T}}$ and regressor matrix $H = (H_{10}, \ldots, H_{1M}, \ldots, H_{p0}, \ldots, H_{pM})$ where $H_{k\ell} = (\phi_\ell(p+1)x_{p+1-k}, \ldots, \phi_\ell(T)x_{T-k})^{\mathrm{T}}$ and $(\cdot)^{\mathrm{T}}$ denotes transpose [13]. The least squares problem can now be stated as

$$\min_C \|X - HC\|_2 \ , \tag{3}$$

the formal solution of which is

$$C = (H^{\mathrm{T}}H)^{-1}H^{\mathrm{T}}X \ . \tag{4}$$

Numerical problems may occur, if the process is almost predictable and the model order p is selected too large. The time-varying coefficients are then assembled via (2). An example of ERD and the estimated coefficients of TVAR(4) model are shown in Fig. 1.

A method that considers catenated coefficient evolutions and their estimation was described in [13, 18]. The reformulation draws the coefficient transition times to occur simultaneously but has the drawbacks that were mentioned in Section 1.

2.2 Selection of basis functions for transition dynamics

Several sets of functions have been used with the TVARLS model and the selection of basis functions has been discussed widely in the literature, see for example [13, 27]. Basis functions that are

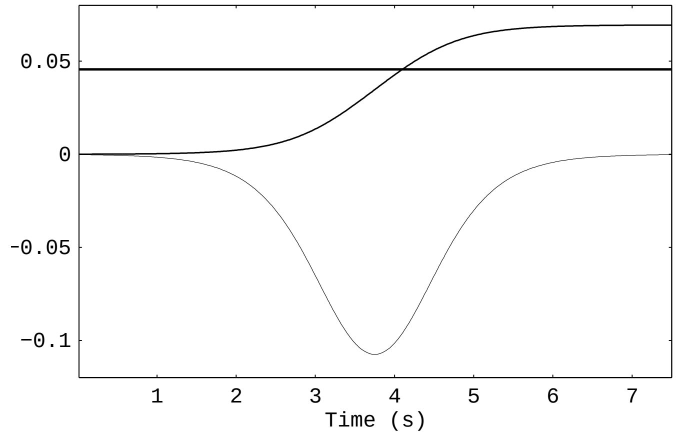

Figure 2: The normalized constant vector (bold line) and the first (medium line) and second (weak line) eigenvectors of the covariance of the set of sigmoids, $\phi_\ell(t)$, $\ell = 0, 1, 2$, respectively.

commonly used to model smooth changes include polynomial, Fourier and discrete spheroidal bases.

However, with these generic bases the parameter estimation problem has too many degrees of freedom in the case of the ERD/ERS transition. In this case we can usually assume that the dynamics can be described as a transition from a state to another. With this assumption we can tailor the basis functions so that they allow changes only in the assumed transition region. Such a set of basis functions can be constructed as in [13] in which a set of sigmoidal transition models for the coefficients was constructed and an optimal low-dimensional approximating set was determined. In this method, however, the coefficient evolutions were catenated.

We use this approach in the construction of (some of) the basis functions also in this paper. A set of sigmoidal functions with smooth transitions from 0 to 1 are first constructed. These functions form a subspace of $\mathbb{R}^T$. The optimal (in the 2-norm sense) two dimensional approximating subspace is then constructed. These two basis functions together with the constant function are used as the basis of the proposed method. The three basis functions are shown in Fig. 2.

Consider the time-varying coefficient $a_k(t)$. Let the coefficient c_{k1} be positive. It can then clearly be seen that if c_{k2} is also positive, the transition occurs earlier than in the case in which c_{k2} is negative. On the other hand, if c_{k1} and c_{k2} are both negative, the transition occurs also early, see also Fig. 3. In addition to being structurally able to model the instant of a transition with a few basis functions, this implies, that a suitable constraint could be used to force the transitions of each of the coefficient evolutions to occur at least approximately simultaneously.

2.3 The β–parametrization of the TVARLS problem

Our aim is to reparametrize the TVARLS problem so that the transition instant can be modelled with a single parameter. We have now $M = 2$ so that the case $c_{k2} = 0$ corresponds to a situation that the transition occurs in the middle of the interval, see Fig. 3.

We modify the original problem (3) problem

$$\min_C \|HC - X\|, \tag{5}$$

by applying the constraints

$$c_{k2} = \beta c_{k1}, \quad \text{for all } k = 1, \ldots, p\ . \tag{6}$$

We have thus $C = (c_{10}, c_{11}, \beta c_{11}, \ldots, c_{p0}, c_{p1}, \beta c_{p1})^{\mathrm{T}}$ and it is easy to show that the transition instants of each coefficient evolution coincide exactly. Thus the transition instants are constrained

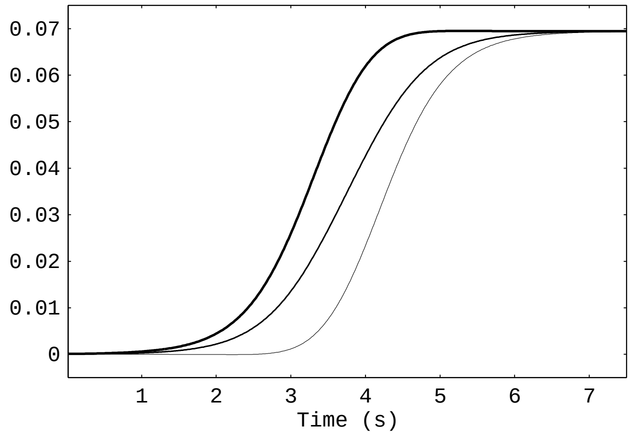

Figure 3: The dependence of a coefficient evolution $a_k(t) = \sum_{\ell=0}^{2} c_{k\ell}\phi_\ell(t)$ of the parameter β. Here $c_{k0} = 0,\ c_{k1} = 1$ and $c_{k2} = \beta c_{k1}$. Bold line: $\beta = 0.2$, medium line: $\beta = 0$ and weak line: $\beta = -0.2$.

to occur simultaneously for all coefficients and this instant depends only on the single parameter β. The dependence of the transition can be seen in Fig. 3. The smaller the β, the later the transition regardless whether the parameters c_{1k} are negative or positive.

We expand the term HC to obtain

$$\begin{aligned} HC &= H_{10}c_{10} + H_{11}c_{11} + H_{12}\beta c_{11} + \cdots + \\ &\qquad H_{p0}c_{p0} + H_{p1}c_{p1} + H_{p2}\beta c_{p1} \qquad (7) \\ &= \sum_{k=1}^{p} (H_{k0}c_{k0} + H_{k1}c_{k1} + \beta H_{k2}c_{k1}) \qquad (8) \\ &\doteq F(\alpha)\ , \qquad (9) \end{aligned}$$

where $\alpha = (c_{10}, c_{11}, \ldots, c_{p0}, c_{p1}, \beta)^{\mathrm{T}}$. This mapping $\alpha \mapsto F(\alpha)$ is nonlinear and thus the obtained least squares problem

$$\min_{\alpha} \|F(\alpha) - X\| \qquad (10)$$

is nonquadratic and has to be solved iteratively. We solve this problem with the Levenberg-Marquardt method (stabilized Gauss-Newton algorithm). We refer to the solutions of the TVARLS problem with parameters α as TVARLS-β estimates in the sequel.

The Jacobian $J_F(\alpha)$ of $F(\alpha)$ is obtained by direct calculation and is given columnwise by

$$\begin{aligned} J_F(\alpha) &= \Big(H_{10}, (H_{11} + \beta H_{12}), \ldots, H_{p0}, (H_{p1} + \beta H_{p2}), \\ &\qquad \sum_{k=1}^{p} H_{k2}c_{k1} \Big) \in \mathbb{R}^{T-p \times 2p+1}\ . \qquad (11) \end{aligned}$$

The Levenberg-Marquardt algorithm takes the form [28]

$$\begin{aligned} \alpha^{(\ell+1)} &= \alpha^{(\ell)} + \mathrm{step}^{(\ell)} \left(J_F^{\mathrm{T}}(\alpha^{(\ell)}) J_F(\alpha^{(\ell)}) + \gamma I \right)^{-1} J_F^{\mathrm{T}}(\alpha^{(\ell)}) \\ &\qquad \left(X - F(\alpha^{(\ell)}) \right)\ , \qquad (12) \end{aligned}$$

where $\gamma > 0$ is a (small) stabilization parameter, I is the identity matrix and $\mathrm{step}^{(\ell)}$ is the step size. Normally $10 - 20$ iterations are adequate to achieve sufficient accuracy if $\mathrm{step}^{(\ell)}$ is set to a constant. If one-dimensional search is used to adjust step$^{(\ell)}$, the stability and in most cases also the speed of the algorithm will be enhanced.

The stabilization parameter γ is needed for the cases in which the Jacobian J_F is ill-conditioned. This may occur in principle at any time during the iteration even if the Jacobian would be well-conditioned at the solution α_* for which the orthogonality criterion $J_F^{\mathrm{T}}(X - F(\alpha_*)$ holds. The solution is not affected by γ but the convergence rate depends this parameter. A large γ protects from erroneous large norm errors in the direction estimates in the Levenberg-Marquardt algorithm which is especially relevant when one-dimensional search is not implemented. On the other hand, a large γ reduces the quadratic nature of the convergence of the Gauss-Newton and weakly stabilized Levenberg-Marquardt algorithms to linear convergence.

The ill-conditionality of J_F can occur easily when the synchronized state is almost predictable, that is, the variance of the so-called innovations of the process (in the ideal case this coincides with the residuals) is small compared to the variance of the process itself [29]. Then, if the model order p is greater than the dimension of the minimal realization of the optimal predictor, the Jacobian J_F and thus also $J_F^{\mathrm{T}} J_F$ are almost singular. This results in a heavily erroneous search direction $(J_F^{\mathrm{T}} J_F + \gamma I)^{-1} J_F^{\mathrm{T}}(X - F(\alpha))$ the norm of which can be very large if γ is small.

Another source of possible instability is the following. Assume that there is no change in the process statistics. The estimator will the yield small estimates for all coefficients c_{1k} (ideally $c_{k1} = 0$ for al k). But we would then also have $c_{k2} \approx 0$ and β could be arbitrary. This in turn means that the rank of the Jacobian is $2p$ instead of $2p+1$. There are many approaches to solve this problem, such as a treshold that is compared to $\max_k |c_{k1}|$ to determine whether there actually was no change in the statistics. A more systematic approach would be the following. First form stationary models for the initial and final (synchronized and desynchronized) states. Then use a test that is constructed to give an estimate for the probability that the parameter differences in these models are due to statistics only.

3 Experimental results

3.1 Simulations and the transition instant estimator $\widehat{T}_{\mathrm{d}} = \widehat{T}_{\mathrm{d}}(\beta)$

We constructed a set of 14 different time-varying AR(6) processes as in [30]. These processes simulate ERD transitions that occur smoothly. The dynamics of the simulations is realized with Gaussian-shaped basis and is thus different from the basis that is used in the estimation. For each of the processes 10 realizations at 11 different transition instants were simulated resulting in 1540 cases of ERD. The β parameter was estimated for each of these ERD simulations, see Fig. 4 for examples of the simulations. These ERD simulations that all have exatly the same statistical time-varying properties indicate clearly the problem that is associated with the transition estimation.

In principle the dependence between β and the transition time could be obtained directly by solving the equation

$$\phi_1(T_{\mathrm{d}}) + \beta\phi_2(T_{\mathrm{d}}) = 1/2\phi_1(T)$$

for each β. However, the resulting estimator $\widehat{T}_{\mathrm{d}} = \widehat{T}_{\mathrm{d}}(\beta)$ is biased in practice. For this reason we aim for such an estimator in which β is based on real TVARLS estimates for such simulations that have realistic transient characteristics and known (by construction) transition instants. Thus the bias that is associated with the estimation of the parameter β will be at least partially avoided in the estimates $\widehat{T}_{\mathrm{d}}(\beta)$.

We seek a linear estimate for the dependence between the transition instants and the β estimates

$$\widehat{T}_{\mathrm{d}} = \psi_1 + \psi_2\beta. \tag{13}$$

To obtain an estimate for a feasible $\Psi = (\psi_1, \psi_2)^{\mathrm{T}}$ we solve the mixed norm problem

$$\min_{\Psi} \left\{ \sum_{jk\ell} |T_{\mathrm{d}}(j) - \psi_1 - \psi_2\beta_{jk\ell}|^1 + \epsilon \sum_{j} |T_{\mathrm{d}}(j) - \psi_1 - \psi_2\bar{\beta}_j|^2 \right\}, \tag{14}$$

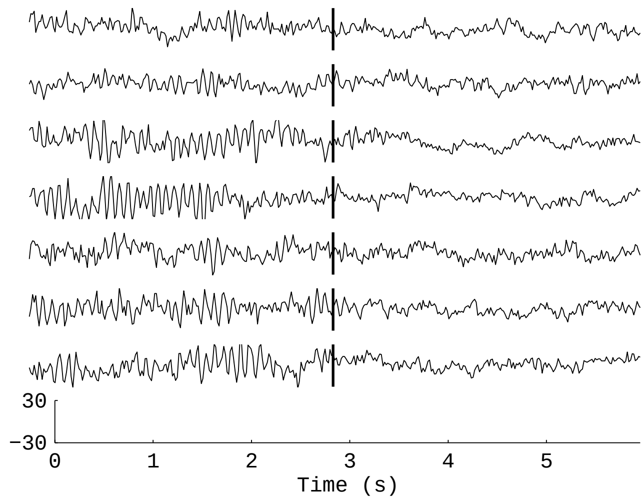

Figure 4: Seven examples of simulated time-varying EEG (ERD). The evolution of coefficients in this model is the same for all simulations. The vertical lines indicate the center of the transition region. The vertical scale is in microvolts.

where ϵ is a weighting parameter between the 1-norm and 2-norm, $\beta_{jk\ell}$ is the estimate for the j'th delay of the k'th realization of the ℓ'th process and $\bar{\beta}_j$ is the mean over all realizations and processes with delay $T_{\rm d}(j)$. The selection $\epsilon = 1000$ turned out to be a feasible weight.

The minimization of (14) seeks to achieve a balance between the delay errors $T_{\rm d} - \widehat{T}_{\rm d}$ for each of the 1540 cases and the mean delay estimate $\widehat{T}_{\rm d}$ over all realizations, see Fig. 5. The one–norm minimization of individual delays is robust and aims to prevent possibly unstable β estimates from inducing large errors to the Ψ estimate. The mean delay estimate is shown in Fig. 5 and an example of the delay tracking characteristics is shown in Fig. 6.

The duration of the transition period of the true coefficient evolutions is approximately 800 ms. Examples of the tracking of the transitions are shown in Fig. 7. The mean error for the transitions is less than one third of the duration of the transition. Thus the proposed method is able to distinguish the (predefined) transition instant even within the transition period. The estimation of this instant within a set of realizations from a particular process is problematic for segmentation algorithms.

3.2 Estimation of event–related desynchronization

We analyze ERD data of 22 patients in order to evaluate the estimation capabilities of the proposed method. The ERD data were measured from occipital region, that is, channels O1 and O2 in the 10/20 electrode placement system. The measurement situation was the basic ERD/ERS test in which the patient opened and closed her/his eyes according to an auditory stimulus. The sampling rate was 256 Hz but the samples were low–pass filtered and decimated to decrease the sampling rate to 256/4 Hz.

In contrast to the simulation study in which we can try to adjust Ψ so that the transition estimate $\widehat{T}_{\rm d}$ really estimates, for example, the center of the transition region, with real ERD data we do not know the true transition instants. However, the true (defined) transition instant affects only the parameter ψ_1 and it has no effect to the parameter ψ_2. Thus we can take a single sample, construct a number of different delays of this sample and assign these samples virtual transition times $T_{\rm d}$ that equal these delays.

In the sequel the term sample refers to a measurement vector the center of which coincides loosely with the auditory stimulus (trigger). Further, the term learning set is used to refer to those samples that were used in the determination of the regressor vector Ψ. These samples are not used

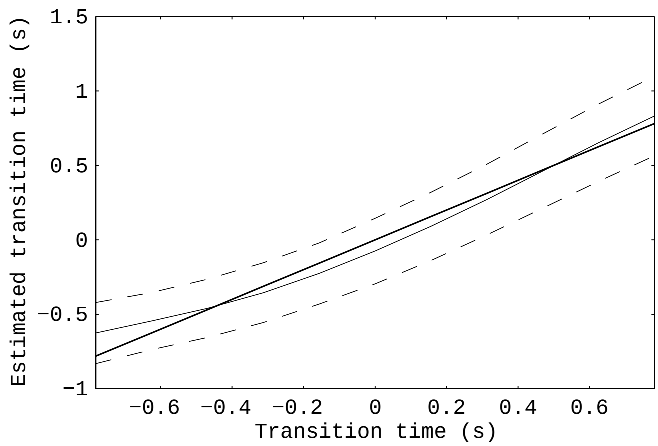

Figure 5: The true delay T_d (bold), the mean of the estimated delays $\widehat{T}_\mathrm{d}$ (weak) and the mean error intervals.

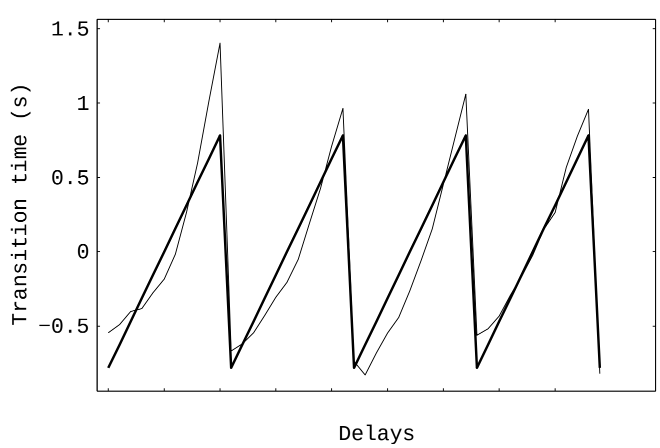

Figure 6: The true delays T_d (bold) and the estimated delays $\widehat{T}_\mathrm{d} = \psi_1 + \psi_2\widehat{\beta}$ (weak) for a fixed evolution (process), 11 delays and 4 realizations shown consecutively. Each ramp corresponds to a single realization with different delays (transitions with respect to time).

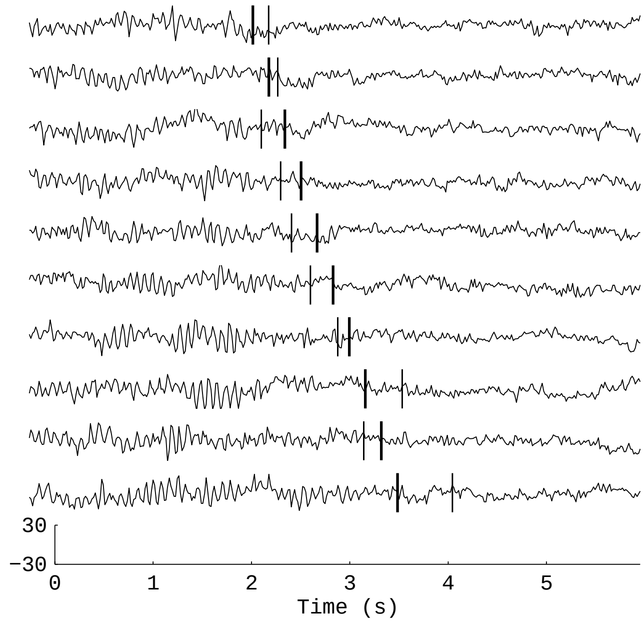

Figure 7: Delay estimates for a process with different delays, true delays T_{d} (bold) and the estimates $\widehat{T}_{\mathrm{d}}$ (weak). The vertical scale is in microvolts.

in the evaluation of the estimation.

We study the ERD samples and transition estimates of three individual patients and one group of patients. We performed the following steps in the estimation of the transition instants.

1. A group of patients (or one patient) were chosen for the estimation.
2. A number of $22-26$ samples per patient were recorded. Of these samples 12 were used for the learning set. Both occipital channels (O1 and O2) were measured (11–13 from each channel).
3. Eleven delayed version of each sample were constructed. The length of the samples was set to $T=400$ that corresponds to 6.25 seconds.
4. The TVARLS-β estimates for each sample of learning set were calculated.
5. The mixed norm problem (14) is solved to obtain the parameters $\Psi=(\psi_1,\psi_2)$.
6. The parameters β for each ERD sample of test set are solved and the estimated transitions $\widehat{T_d}$ in (13) are calculated.

The results are shown in Figs. 8, 9, 10 and 11. The bold, vertical lines indicate the time instant when the patients heard the auditory stimulus and opened their eyes (stimulus time). The weak lines indicate the transition estimates that were obtained using the proposed method. All the shown ERD samples are from the test sets.

Case 1: Strongly modulating ERD, Fig. 8. These kind of samples are sometimes problematic to segmentation algorithms that are not based on the prediction error residuals. This is an easy case that could relatively well be segmented also visually. The proposed method is also able to track the transitions and the estimates are reasonable. However, if all ERD/ERS cases were like this, there would be little need for such algorithms as the one considered in this paper.

Case 2: Smooth transitions, Figs. 9 and 10. The transitions are unclear in these cases and there are hardly any visible differences in the transition area. The alpha rhythm vanishes when the eyes

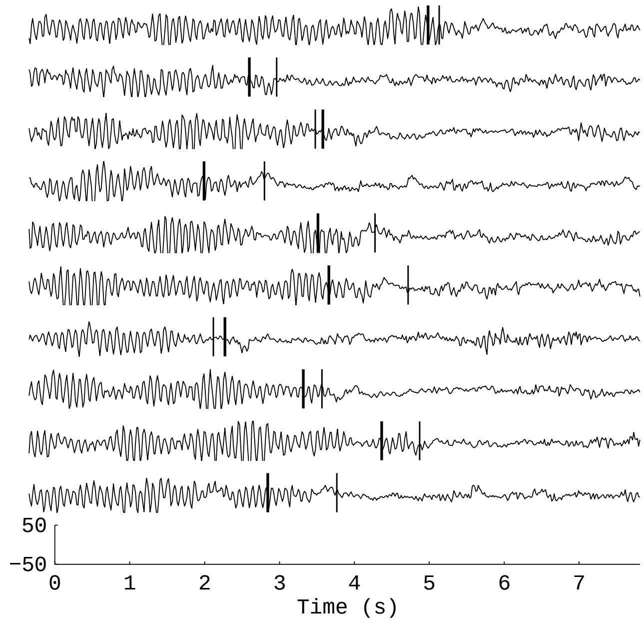

Figure 8: Case 1. ERD samples of a healthy, young male. Stimulus times (bold line) and transition estimates (weak line). The vertical scale is in microvolts.

are opened but it reappears in the eyes open state. This kind of event–related desynchronization causes problems to most segmentation methods and it would be excessively difficult to estimate the transition instants visually. The proposed method, however, performs relatively well in estimating the transitions.

Case 3: Group of patients, Fig. 11. The ERD samples of three, healthy persons were used both for the learning set and the test set. All samples exhibit from moderate to heavy modulation with more or less unclear transitions. The proposed method is able to track the transitions reasonably well. In this example it was possible to construct the learning set with data from different test persons.

As is shown, the method is able to track smooth medium rate transitions also in such cases that are problematic for visual change instant detection and segmentation algorithms. The ultimate goal of such an estimation method that is described here, is that the learning phase should be done only once, that is, a single vector Ψ (in fact only ψ_1) could be used with all ERD samples and the estimated β parameters. However, the bias in the estimation of the β parameter seems to depend on the synchronized and desynchronized states so much that this end is not achieved with the TVARLS-β algorithm without further development.

4 Discussion

We have proposed a method for the modelling of smooth medium rate transitions of EEG such as event–related synchronization/desynhronization. The method is is a modification of a more general TVARLS scheme that constraints the transitions of all coefficient evolutions to occur simultaneously. The method can be used in such cases in which the instant of transition is only approximately known, which is the case in ERD/ERS. The more general TVARLS model that can describe more complicated variations in EEG transitions, such as the rate of transition and still maintain the desired simultaneity of coefficient transitions, are much more complex and are feasible in the analysis of a single patient only.

The performance of the method was evaluated with simulations of known transition dynamics. The method was shown to be able to track the instant of a smooth medium rate transition. Such

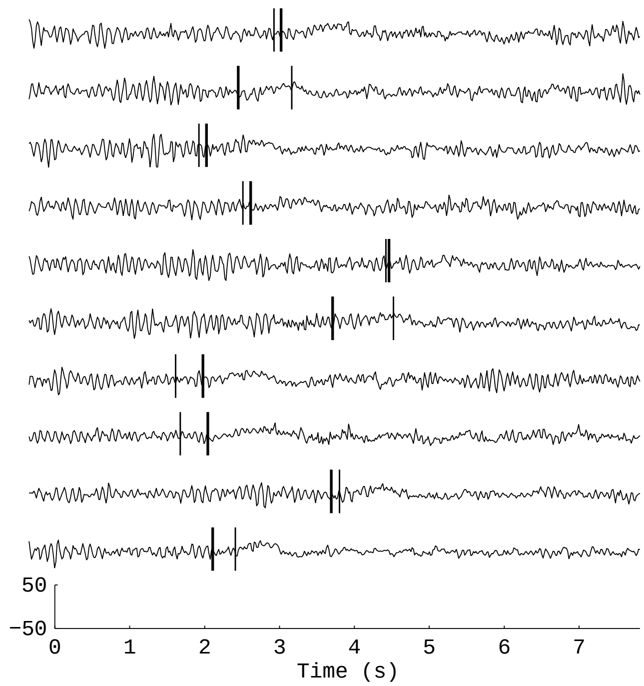

Figure 9: Case 2. ERD samples of an Alzheimer patient. Stimulus times (bold line) and transition estimates (weak line). The vertical scale is in microvolts.

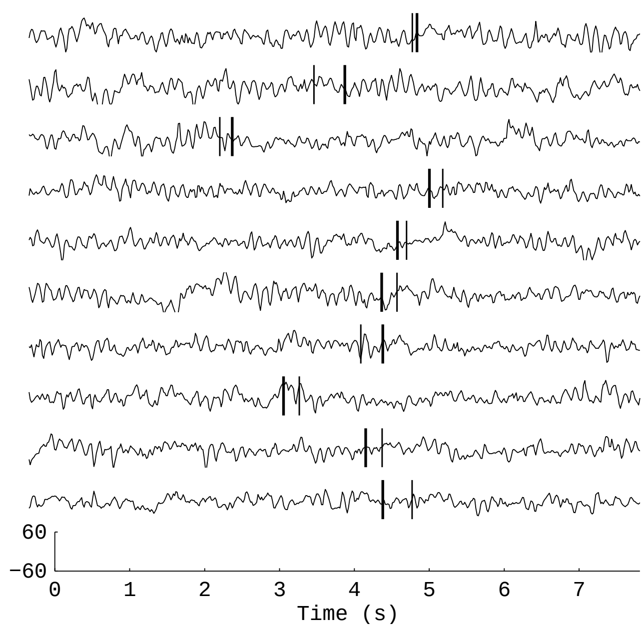

Figure 10: ERD samples of a multi–infarct patient. Stimulus times (bold line) and transition estimates (weak line). The vertical scale is in microvolts.

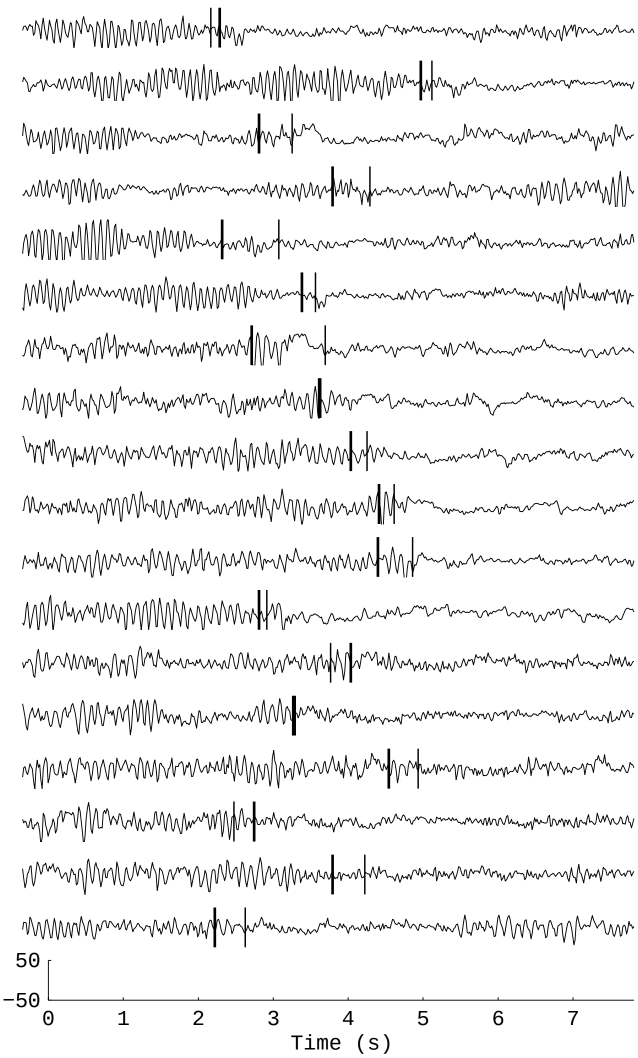

Figure 11: Case 3. ERD samples of three, healthy persons, 6 samples from each person. Stimulus times (bold line) and transition estimates (weak line). The vertical scale is in microvolts.

transitions conflict with the assumptions that are inherent when conventional segmenters and, on the other hand, adaptive algorithms are used.

Finally, the method was applied to event–related desynchronization data of healthy persons as well as patients who have Alzheimers disease or multi–infarct dementia (MID). The method was able to track the transitions in such cases that are problematic to conventional approaches. This suggests that the proposed TVARLS-β algorithm, whose assumptions on the rate of change in the process statistics fall between the segmentation algorithms and adaptive algorithms, can be further developed so that the estimator for the transition instant would not depend notably on the synchronized and desynchronized states. The most obvious but also most demanding approach to achieve this end is to estimate the conditional distribution of the transition instant on all α parameters. This approach, that is Bayesian in nature, is also structurally able to take into account the bias in β that depends on the synchronized and desynchronized states.

Since the applications for which the proposed method is intended are basically the same as those in which segmenters are usually used with, it is necessary to compare the performances of TVARLS-β and segmenters. However, it has turned out that the detection capabilities of change points are very different with different segmenters. On the other hand, it is relatively easy to construct such simulations that comply perfectly with a particular method. The results will then naturally be superior with this method when compared to those methods that are based on different (implicit) assumptions. Furthermore, while the proposed method always suggests a transition instant, a segmenter might not do that if the differences in the statistics are small enough. A fair comparison is therefore a difficult task that is left for further studies. However, if a rate of transition can be established with some method, the proposed TVARLS-β scheme can easily be adjusted to conform to this rate.

References

[1] J. Barlow, "Methods of analysis of nonstationary EEGs, with emphasis on segmentation techniques: A comparative review," *Electroencephalogr Clin Neurophysiol*, vol. 2, no. 3, pp. 267–304, 1985.

[2] J. Barlow, O. Creutzfeldt, D. Michael, J. Houchin, and H. Epelbaum, "Automatic adaptive segmentation of clinical EEGs," *Electroencephalogr Clin Neurophysiol*, vol. 51, pp. 512–525, 1981.

[3] D. Michael and J. Houchin, "Automatic EEG analysis: A segmentation procedure based on the autocorrelation function," *Electroencephalogr Clin Neurophysiol*, vol. 46, pp. 232–235, 1979.

[4] G. Bodenstein and H. Praetorius, "Feature extraction from the electroencephalogram by adaptive segmentation," *Proc IEEE*, vol. 65, pp. 642–652, 1977.

[5] G. Bodenstein and H. Praetorius, "Pattern recognition of EEG by adaptive segmentation," in *Biomedical Computing* (W. Perkins, ed.), pp. 20–31, Pitman Medical, 1977.

[6] G. Bodenstein, W. Schneider, and C. Malsburg, "Computerized EEG pattern classification by adaptive segmentation and probability-density-function classification. Description of the method," *Comput Biol Med*, vol. 15, pp. 297–313, 1985.

[7] T. Bohlin, "Analysis of EEG signals with changing spectra using a short-word Kalman estimator," *Math Biosci*, vol. 35, pp. 221–259, 1977.

[8] S. Cerutti, D. Liberati, G. Avanzini, S. Franceschetti, and F. Panzina, "Classification of the EEG during neurosurgery. Parametric identification and Kalman filtering compared," *J Biomed Eng*, vol. 8, pp. 244–254, 1986.

[9] G. Müller, R. A. Richter, S. Weisbrod, and F. Klingberg, "Duration of EEG alpha wave blockade by tone stimulation is prolonged in early stage of presenile onset dementia of the Alzheimer type," *Biomed Biochim Acta*, vol. 50, no. 8, pp. 987–991, 1991.

[10] L. Defebvre, P. Derambure, J. L. Bourriez, J. M. Jacquesson, K. Dujardin, A. Destée, and J. D. Guieu, "Spatiotemporal study of event–related desynchronization in idiopathic Parkinson's disease," in *Advances in Neurology* (H. Narabayashi, T. Nagatsu, N. Yanagisawa, and Y. Mizuno, eds.), (New York), pp. 422–428, Raven Press, 1993.

[11] L. Defebvre, J.-L. Bourriez, K. Dujardin, P. Derambure, A. Destée, and J.-D. Guieu, "Spatiotemporal study of bereitschaftspotential and event–related desynchronization during voluntary movement in Parkinson's disease," *Brain Topography*, vol. 6, no. 3, pp. 237–244, 1994.

[12] C. Grillon and M. Buchsbaum, "EEG topography of response to visual stimuli in generalized anxiety disorder," *Electroencephalogr Clin Neurophysiol*, vol. 66, pp. 337–347, 1987.

[13] J. Kaipio, *Simulation and Estimation of Nonstationary EEG*. PhD thesis, University of Kuopio, 1996.

[14] G. Pfurtscheller, "Event–related synchronization (ERS): an electrophysiological correlate of cortical areas at rest," *Electroencephalogr Clin Neurophysiol*, vol. 83, pp. 62–69, 1992.

[15] G. Pfurtscheller, C. Neuper, and W. Mohl, "Event-related desynchronization (ERD) during visual processing," *Int J Psychophysiol*, vol. 16, pp. 147–153, 1994.

[16] O. Markand, "Alpha rhythms," *J Clin Neurophysiol*, vol. 7, pp. 163–189, 1990.

[17] E. Niedermeyer and F. L. da Silva, eds., *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Williams and Wilkins, third ed., 1993.

[18] J. Kaipio and P. Karjalainen, "Estimation of event related synchronization changes by a new TVAR method," *IEEE Trans Biomed Eng*, vol. 44, no. 8, 1997.

[19] J. Hiltunen, P. Karjalainen, and J. Kaipio, *Estimation of Medium Rate EEG Transitions*. EMBS97 Conference, 1997.

[20] T. S. Rao, "The fitting of non-stationary time series models with time-dependent parameters," *J R Statist Soc*, vol. B 32, pp. 312–322, 1970.

[21] L. Liporace, "Linear estimation of nonstationary signals," *J Acoust Soc Am*, vol. 58, pp. 1288–1295, 1975.

[22] W. Gersch, J. Yonemoto, and P. Naitoh, "Automatic classification of multivariate EEGs using an amount of information measure and the eigenvalues of parametric time series model features," *Comput Biomed Res*, vol. 10, pp. 297–318, 1977.

[23] W. Gersch and J. Yonemoto, "Parametric time series models for multivariate EEG analysis," *Comput Biomed Res*, vol. 10, pp. 113–125, 1977.

[24] M. Hall, A. Oppenheimer, and A. Willsky, "Time-varying parametric modeling of speech," *Signal Processing*, vol. 5, pp. 267–285, 1983.

[25] Y. Grenier, "Time-dependent ARMA modeling of nonstationary signals," *IEEE Trans Acoust, Speech Signal Processing*, vol. ASSP-31, pp. 899–911, 1983.

[26] R. Charbonnier, M. Barlaud, G. Alengrin, and J. Menez, "Results on AR-modelling of non-stationary signals," *Signal Processing*, vol. 12, pp. 143–151, 1987.

[27] N. Amir and I. Gath, "Segmentation of EEG during sleep using time varying autoregressive modeling," *Biol Cybern*, vol. 61, pp. 447–455, 1989.

[28] C. Lawson and R. Hanson, *Solving Least Squares Problems*. SIAM, 1995.

[29] A. Papoulis, *Probability, Random Variables and Stochastic Processes*. McGraw-Hill, 1984.

[30] J. Kaipio and P. Karjalainen, "Simulation of nonstationary EEG," *Biol Cybern*, vol. 76, 1997.