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

# Estimation of Nonstationary EEG with Kalman Smoother Approach: an Application to Event-Related Synchronization (ERS)

M.P. Tarvainen, J.K. Hiltunen,
P.O. Ranta-aho, and P.A. Karjalainen

June 19, 2003

Report No. 1/2003

This manuscript has been submitted to *IEEE Trans Biomed Eng*

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

M.P. Tarvainen*, J.K. Hiltunen†, P.O. Ranta-aho, and P.A. Karjalainen

June 19, 2003

**Abstract** An adaptive spectrum estimation method for nonstationary EEG by means of time-varying autoregressive moving average modeling is presented. The time-varying parameter estimation problem is solved by Kalman filtering along with a fixed-interval smoothing procedure. Kalman filter is an optimal filter in the mean square sense and it is a generalization of other adaptive filters such as recursive least squares (RLS) or least mean square (LMS). Furthermore, by using the smoother the unavoidable tracking lag of adaptive filters can be avoided. Due to the properties of Kalman filter and benefits of the smoothing the time-frequency resolution of the presented Kalman smoother spectra is extremely high. The presented approach is applied to estimation of event-related synchronization/desynchronization (ERS/ERD) dynamics of occipital alpha rhythm measured from three healthy subjects. With the Kalman smoother approach detailed spectral information can be extracted from single ERS/ERD samples.

Kalman filter, smoothing, spectrogram, event-related synchronization (ERS), event-related desynchronization (ERD), alpha rhythm

## 1 Introduction

The electroencephalogram (EEG) recording is a useful tool for studying the functional state of the brain and for diagnosing certain neurophysiological states and disorders. EEG signals are often quantified based on their frequency-domain characteristics. Typically the spectrum is estimated using the fast Fourier transform (FFT). A fundamental requirement in the FFT-based spectral analysis is the stationarity of the analyzed signal. In [1] it was suggested that EEG epochs shorter than 12 seconds may be considered stationary. However, it is well known that EEG can exhibit considerable short-term nonstationarities. In such situations, time-frequency representation (TFR) methods are required.

A time-frequency representation describes the energy density of the signal simultaneously in time and frequency. A traditional TFR method is the moving window Fourier transform which is also known as spectrogram. In this method the signal is implicitly assumed to be stationary within each windowed segment and the selection of the length of this window is followed by a trade-off between time and frequency resolutions. Another popular TFR method is the Wavelet transform in which the window width is inversely proportional to the frequency resulting in improved time resolution at high frequencies and frequency resolution at low frequencies. In addition to these two linear TFRs various quadratic TFRs such as the Wigner distribution and its numerous smoothed versions have been proposed. For good tutorials on time-frequency representations see e.g. [2, 3, 4].

Time-frequency representations have various applications in the field of EEG analysis. One of the most common applications is the analysis of event-related EEG synchronization (ERS) and desynchronization (ERD) [5, 6, 7, 8]. The terms ERS and ERD usually refer to stimulus-induced changes in EEG amplitude or power within certain frequency band rather than synchrony. This is also the case in this paper and, therefore, the present study should not be confused to studies of EEG synchrony such as [9]. Furthermore, the ERS/ERD changes are time but not phase-locked to the entailed event, and cannot, therefore, be extracted by simple linear methods such as averaging, but may be detected by frequency analysis [10]. Traditionally the ERD or ERS in

*M.P. Tarvainen, P.O. Ranta-aho, and P.A. Karjalainen are with the Department of Applied Physics, University of Kuopio, P.O.Box 1627, FIN-70211 Kuopio, Finland

†J.K. Hiltunen is with the Section of Clinical Neurosciences, Finnish Institute of Occupational Health, Helsinki, Finland.

a specified frequency band is quantified by bandpass filtering, squaring of samples, and averaging over trials [11]. A drawback of this approach is the rather low frequency resolution. Furthermore, the selection of the band limits can be problematic, even though, the limits are nowadays usually adjusted individually based on some specific frequency. For a review on ERD estimation methods see e.g. [12]. TFR methods have also been used in the inspection of the frequency content of event-related brain potentials [13, 14], quantification of epileptic seizure dynamics [15, 16, 17, 18], and analysis of newborn EEG seizure events [19, 20].

Nowadays, one of the most popular TFR methods used in EEG analysis, and in biomedical applications in general, is the Wavelet transform [21, 22]. This method, however, suffers from the same kind of trade-off between time and frequency resolutions as the traditional spectrogram method. An improved time-frequency resolution can be obtained by using parametric spectral analysis methods based on time-varying linear models. A common approach is to use a time-varying autoregressive moving average (ARMA) model. The frequency resolution of parametric methods is superior because of the implicit extrapolation of the autocorrelation sequence [23]. For the same reason the leakage effect of the classical spectral estimators, depending on the used windowing function, is suppressed.

The main task in parametric modeling is the estimation of the time-varying model parameters. For this, adaptive algorithms such as Kalman filter, least mean square (LMS), and recursive least squares (RLS) have been adopted. The Kalman filter which was originally introduced in [24, 25] is an optimal estimator for time-varying parameters in mean square sense and is based on so called state-space formalism. Kalman filter has been previously used in EEG analysis in e.g. [26, 27, 28, 29]. A good overview on the analysis of nonstationary EEG with parametric models can be found in e.g. [30, 31]

A drawback of the Kalman filter, as of all other adaptive algorithms, is the tracking lag present in the estimated parameters. This is especially disadvantageous when the aim is to estimate accurately some abrupt (perhaps event-related) changes of EEG. However, the tracking lag can be avoided by using the so-called smoother algorithm with the Kalman filter. A smoother is an estimator which utilizes the future measurements in addition to the past ones when computing the estimates at a given time point. Three different types of smoothers, namely fixed-point, fixed-lag, and fixed-interval smoothers, have been proposed e.g. in [32, 33], from which an appropriate one can be selected. An estimator including the Kalman filter along with a smoother is called Kalman smoother.

In this paper we present a Kalman smoother approach, utilizing the fixed-interval smoother, for estimating time-frequency structures of nonstationary EEG signals. The tracking ability of the presented approach is tested with simulations and the advantages against commonly used adaptive filters are pointed out by comparing Kalman smoother with the popular forgetting factor RLS algorithm. As a specific application Kalman smoother is applied to quantification of ERS/ERD dynamics of occipital alpha rhythm. With the presented approach detailed time-frequency representations for single ERS/ERD samples can be extracted. Even short-term changes such as the so-called "squeak" effect of alpha rhythm [34] can be observed from single ERS sample spectrum estimates. The term "squeak" is used for the phenomenon of alpha rhythm where the ERS after eye closure starts up at higher frequency and directly begins to shift to lower frequencies. By using the classical spectrogram method such short-term changes can not be observed without averaging several consecutive trials due to poor resolution.

## 2 Methods

In dynamic parametric spectral estimation the measured signal is modeled as an output of a parametric model with time-varying parameters. Model parameters can be estimated with a number of adaptive algorithms such as the Kalman filter which is based on state-space formalism.

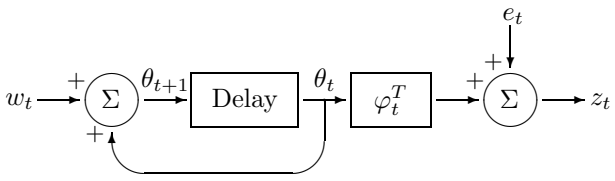


Figure 1: Structure of the state-space signal model.

### 2.1 State-space formulation

Here we use a time-varying autoregressive moving average model for the signal. The time-varying ARMA($p$,$q$) model for signal $z$ at time instant $t$ can be written in the form

$$z_t = -\sum_{j=1}^{p} a_t^{(j)} z_{t-j} + \sum_{k=1}^{q} b_t^{(k)} e_{t-k} + e_t \tag{1}$$

where $a_t^{(j)}$ and $b_t^{(k)}$ are the time-varying AR and MA parameters at time instant $t$. Constants $p$ and $q$ are the corresponding model orders of AR and MA parts and $e_t$ is the observation error. By denoting

$$\theta_t = \left(-a_t^{(1)}, \ldots, -a_t^{(p)}, b_t^{(1)}, \ldots, b_t^{(q)}\right)^T \tag{2}$$

$$\varphi_t = \left(z_{t-1}, \ldots, z_{t-p}, e_{t-1}, \ldots, e_{t-q}\right)^T \tag{3}$$

the time-varying ARMA model can be written in the form

$$z_t = \varphi_t^T \theta_t + e_t. \tag{4}$$

This is formally a linear observation model with $\varphi_t^T$ being the regression vector and $e_t$ the observation error. Note that the sequence $e_t$ in the regression vector is not measured unlike $z_t$, but it has to be estimated together with other parameters.

In the state-space formalism the model parameters $\theta_t$ are also called as states. The variation of the state when no *a priori* information is available is typically described with the random walk model [35]. Thus, we write a state equation

$$\theta_{t+1} = \theta_t + w_t \tag{5}$$

for the parameters. Equations (4) and (5) form a reduced structure of the general state-space equations with the input noise $w_t$. The structure of the signal model is presented in Fig. 1. The task is now to find, in some sense, optimal solution for the time-varying parameters $\theta_t$ according to the above state-space equations.

### 2.2 Kalman filtering

Kalman filter is a tool for estimating the sequence of states of the dynamical system in recursive manner. The formulation of the Kalman filter is described in terms of the state-space concepts. The Kalman filtering problem is to find the minimum mean square estimator $\hat{\theta}_t$ for state $\theta_t$ given the observations $z_1, \ldots, z_t$. In e.g. [36] this has been shown to be equal to conditional expectation

$$\hat{\theta}_t = E\left\{\theta_t | z_1, \ldots, z_t\right\}. \tag{6}$$

The state and measurement noises $w_t$ and $e_t$ are here assumed to be uncorrelated, zero mean, Gaussian white noise processes with covariances $C_{w_t} = \sigma_w^2 I$ and $C_{e_t} = \sigma_e^2$ respectively. Furthermore, the initial state $\theta_0$ is assumed to be a Gaussian random variable with finite variance and uncorrelated with $w_t$ and $e_t$.

TABLE I: Kalman gain vectors $K_t$ and covariance estimates $P_t$ for Kalman filter (KF), RLS, and LMS algorithms

| | $K_t$ | $P_t$ |
|---|---|---|
| KF | $P_{t-1}\varphi_t \left(\varphi_t^T P_{t-1}\varphi_t + C_{e_t}\right)^{-1}$ | $\left(I - K_t\varphi_t^T\right) P_{t-1} + C_{w_t}$ |
| RLS | $P_{t-1}\varphi_t \left(\varphi_t^T P_{t-1}\varphi_t + \lambda\right)^{-1}$ | $\lambda^{-1}\left(I - K_t\varphi_t^T\right) P_{t-1}$ |
| LMS | $\mu\varphi_t$ | $\mu\left(I - \mu\varphi_{t+1}\varphi_{t+1}^T\right)^{-1}$ |

By following the notations used in [37] and [38] the Kalman filter equations can be written in the form

$$\begin{aligned}
\hat{\theta}_{t|t-1} &= \hat{\theta}_{t-1} && (7)\\
C_{\tilde{\theta}_{t|t-1}} &= C_{\tilde{\theta}_{t-1}} + C_{w_{t-1}} && (8)\\
K_t &= C_{\tilde{\theta}_{t|t-1}}\varphi_t \left(\varphi_t^T C_{\tilde{\theta}_{t|t-1}}\varphi_t + C_{e_t}\right)^{-1} && (9)\\
C_{\tilde{\theta}_t} &= \left(I - K_t\varphi_t^T\right) C_{\tilde{\theta}_{t|t-1}} && (10)\\
\epsilon_t &= z_t - \varphi_t^T\hat{\theta}_{t|t-1} && (11)\\
\hat{\theta}_t &= \hat{\theta}_{t|t-1} + K_t\epsilon_t && (12)
\end{aligned}$$

where $\hat{\theta}_{t|t-1}$ is the mean square estimate for state $\theta_t$ given the observations $z_1, \ldots, z_{t-1}$, $\tilde{\theta}_t$ is the state estimation error $\tilde{\theta}_t = \theta_t - \hat{\theta}_t$, and $K_t$ is the Kalman gain vector adjusting the state estimate after each new observation. Note that the unknown observation noise $e_t$ is here estimated with the prediction error process $\epsilon_t$ in every step of the iteration. The adaptation of the filter is primarily affected by $C_{w_t}$ [39].

If we insert (7) and (8) into equations (9)–(12) and denote $P_t = C_{\tilde{\theta}_t} + C_{w_t}$ the Kalman filter recursion can be written in the form

$$\begin{aligned}
\hat{\theta}_t &= \hat{\theta}_{t-1} + K_t\epsilon_t && (13)\\
\epsilon_t &= z_t - \varphi_t^T\hat{\theta}_{t-1} && (14)
\end{aligned}$$

where the Kalman gain vector $K_t$ and the recursive estimate $P_t$ for the covariance $C_{\tilde{\theta}_t} + C_{w_t}$ are

$$\begin{aligned}
K_t &= \frac{P_{t-1}\varphi_t}{\varphi_t^T P_{t-1}\varphi_t + C_{e_t}} && (15)\\
P_t &= \left(I - K_t\varphi_t^T\right) P_{t-1} + C_{w_t}. && (16)
\end{aligned}$$

It turns out that several popular adaptive algorithms can be written in the form (13)–(14) with different choices of $K_t$ and $P_t$. For a detailed description of the relations between RLS and Kalman filter see e.g. [40, 41, 37, 31]. The relations between LMS and Kalman filter have been presented in [37]. The choices of $K_t$ and $P_t$ for Kalman filter, RLS, and LMS are summarized in Table I, where $\lambda$ is the forgetting factor of the RLS (typically $0.9 \leq \lambda \leq 1$) and $\mu$ is the step size of the LMS.

### 2.3 Fixed-interval smoothing

In the Kalman filtering algorithm the state estimate is updated immediately after a new observation is available. However, this kind of processing is not always necessary. Instead, if a small lag in the processing is allowed or if the measured data is not processed in real time also the future observations can be used in the state estimation. In this case it is reasonable to expect the estimates to be more accurate. Such an estimator, which uses observations $z_1, \ldots, z_{t+N}$ to estimate the state $\theta_t$ at time instant $t$, is called a smoother. One such smoother is the fixed-interval smoother.

The fixed-interval smoothing problem is to find the minimum mean square estimator $\hat{\theta}_t$ for each state $\theta_t$ ($t = 1, \ldots, T$) given the observations $z_1, \ldots, z_T$. The smoothed estimate is here denoted as $\hat{\theta}_{t|T}$. As previously, the fixed-interval smoothing problem is equal to the conditional expectation

$$\hat{\theta}_{t|T} = E\left\{\theta_t | z_1, \ldots, z_T\right\} \tag{17}$$

for fixed $T$ and for all $t$ in the interval $1 \leq t \leq T$. The solution for this can be written in the form [32, 33]

$$\hat{\theta}_{t-1|T} = \hat{\theta}_{t-1} + A_{t-1}\left(\hat{\theta}_{t|T} - \hat{\theta}_{t|t-1}\right) \tag{18}$$

$$A_{t-1} = C_{\tilde{\theta}_{t-1}} C^{-1}_{\tilde{\theta}_{t|t-1}} \tag{19}$$

where $A_{t-1}$ includes the error covariances stored in the forward run of Kalman filter. Also the state estimates $\hat{\theta}_t$ and $\hat{\theta}_{t|t-1}$ need to be stored. The smoothed estimates $\hat{\theta}_{t-1|T}$ are then obtained by running the stored estimates backwards in time by taking $t = T, T-1, \ldots, 2$. The initialization is evidently with the filtered estimate, that is $\hat{\theta}_{T|T} = \hat{\theta}_T$.

The fixed-interval smoother is clearly suitable only for off-line processing since the whole measurement set $z_1, \ldots, z_T$ is required in the estimation of each state $\theta_t$ in the interval $t = 1, \ldots, T$. For on-line processing where a small delay between receiving the observation and calculating the corresponding state estimate is allowed the fixed-lag smoother can be used [32, 33].

### 2.4 Spectral estimation

Once the time-varying coefficients of the ARMA($p$,$q$) model (1) are solved the time-varying power spectrum density (PSD) estimate can be obtained in terms of the estimated ARMA coefficients

$$P_t(f) = \sigma_\epsilon^2(t)/f_s \frac{|1 + \sum_{k=1}^{q} b_t^{(k)} e^{-i2\pi k f/f_s}|^2}{|1 + \sum_{j=1}^{p} a_t^{(j)} e^{-i2\pi j f/f_s}|^2} \tag{20}$$

where $\sigma_\epsilon^2(t)$ is the prediction error variance and $f_s$ is the sampling frequency. The PSD estimate can be calculated for each time instant after the adaptive algorithm, used to estimate the time-varying parameters, converges. However, it is not always desirable to calculate the PSD estimate for each time instant, but only within some short time intervals. This can be obtained by averaging the estimated ARMA parameters with a moving window. It is also notable that equation (20) is a continuous function of frequency $f$ and can therefore be evaluated at any desired frequencies from 0 up to the Nyquist frequency $f_s/2$, even though this does not improve the frequency resolution of the spectrum infinitely.

### 2.5 EEG recordings

Event-related synchronization and desynchronization (ERS/ERD) of occipital alpha rhythm was studied from three healthy subjects. In the test procedure subjects closed and opened their eyes in accordance with auditory stimuli at 15 second interstimulus intervals. During the test procedure a total of 19 ERS/ERD samples were obtained for each subject. As a recording device a NeuroScan system (NeuroSoft Inc., Sterling, VA) with the international 10-20 electrode coupling was used. The sampling frequency of the device was 256 Hz. The ERS/ERD for visual stimulation is best seen in the posterior and occipital regions of the brain. Therefore, the occipital O2 channel was selected for the analysis. The sampling frequency of the selected channel was reduced after lowpass filtering to 64 Hz.

## 3 Results

### 3.1 Simulations

The evaluation of performance and estimation accuracy of proposed adaptive algorithm used to model some nonstationary signals is often problematic. The competence of the algorithm, in such

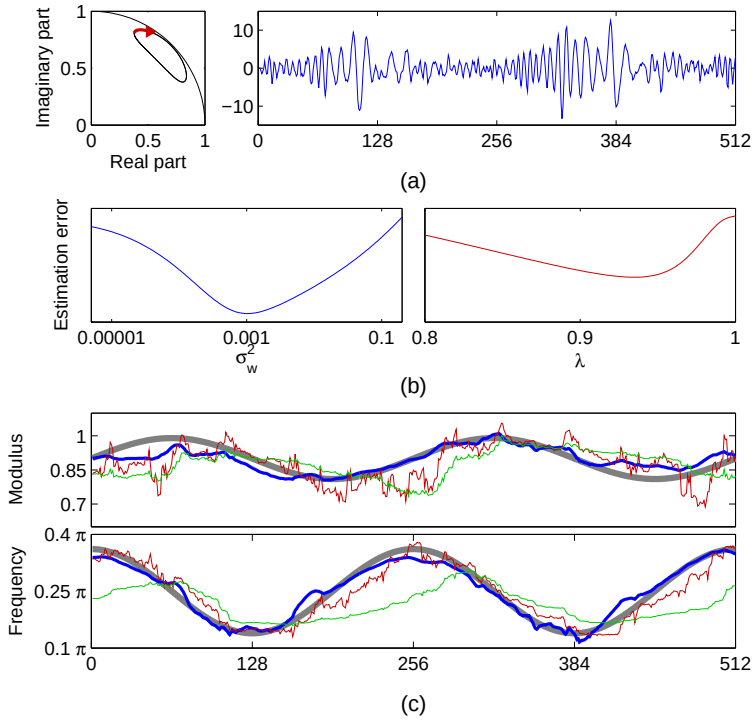


**Figure 2**: AR(2) process estimation with Kalman smoother and RLS algorithms. (a) The root evolution and a typical realization. (b) Optimization of the state noise covariance coefficient ($\sigma_w^2 = 0.001$) of the Kalman filter and of the forgetting factor ($\lambda = 0.935$) of the RLS algorithm. (c) Estimates of the root modulus and frequency: true values (—), Kalman smoother estimates (—), and RLS estimates with $\lambda = 0.935$ (—) and $\lambda = 0.98$ (—).

cases, can be evaluated with simulations. In this paper we conducted two different simulations to test the tracking ability of Kalman smoother. In both simulations Kalman smoother was compared to the popular forgetting factor RLS algorithm.

In the first simulation a smoothly varying AR(2) process was generated. The trace of the simulated AR(2) model root and typical realization are presented in Fig. 2 (a). The parameters controlling the adaptation (i.e. the state noise covariance coefficient $\sigma_w^2$ of Kalman smoother and forgetting factor $\lambda$ of RLS) were selected to minimize the estimation error of AR coefficients [Fig. 2 (b)]. Optimal values for the parameters were $\sigma_w^2 = 0.001$ and $\lambda = 0.935$ and the error for the Kalman smoother was 52 % smaller than for RLS. Obtained estimates for the root modulus and frequency are presented in Fig. 2 (c). The other RLS estimate was calculated with a substantially larger forgetting factor $\lambda = 0.98$. It is clearly seen that by increasing $\lambda$ more stable RLS estimates are obtained but, as a downside, the tracking lag is increased. The tracking accuracy of the Kalman smoother is, however, clearly better than either of the RLS estimates. Unfortunately, not much can be said about the applicability of the Kalman smoother on tracking of nonstationary EEG based on this simple simulation.

In the second simulation an EEG transition from desynchronized to synchronized state was modeled. The used approach is presented in detail in [42], and is, therefore, described here only briefly. For other approaches to simulate nonstationary EEG see e.g. [43, 44]. Occipital EEG recorded with the eyes closed shows high intensity in the alpha frequency band, classically defined as 8–13 Hz band [45]. By opening the eyes this intensity is decreased or even blocked. It can be thought that EEG exhibits a transition from one stationary state to another. Such a transition was here simulated by modeling both stationary states as an fifth-order AR process. The roots of the models were estimated from real EEG measured during an eyes open/closed test. The obtained

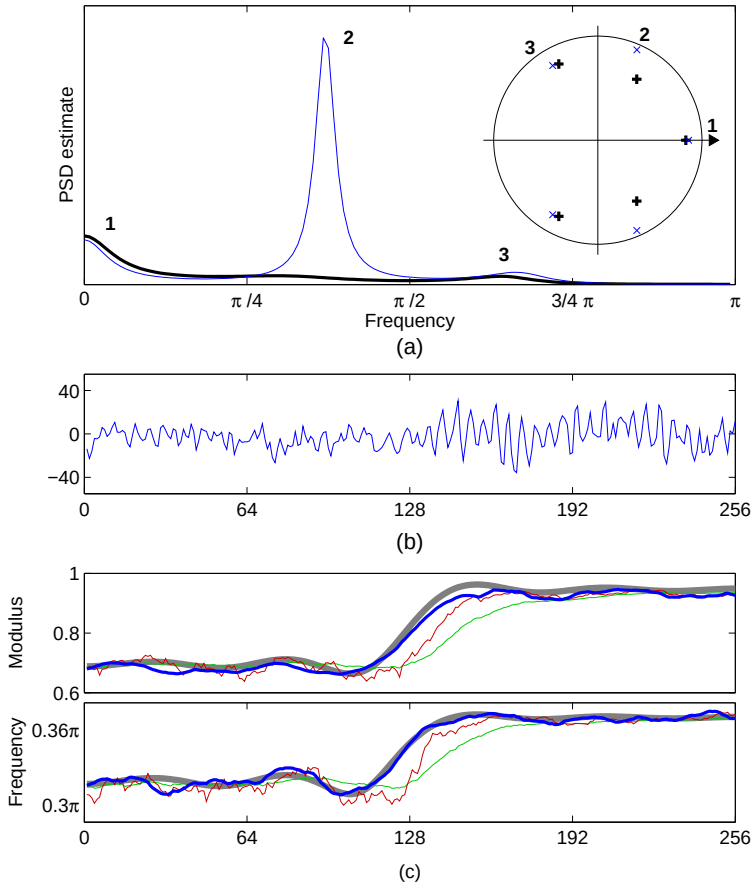


**Figure 3**: A realistic simulation of an EEG transition as an AR(5) process. (a) The roots and the corresponding spectra before (—) and after (—) the transition (eyes open → eyes closed). Numbers 1, 2, and 3 indicate the analogy of the roots and spectral peaks. (b) A typical realization of the smoothed transition. (c) Averaged estimates over 100 realizations of the modulus and frequency of the root 2 corresponding to alpha rhythm: true values (—), Kalman smoother estimates (—) with $\sigma_w^2 = 0.0003$, and RLS estimates with $\lambda = 0.9$ (—) and $\lambda = 0.98$ (—).

roots for both states and the corresponding power spectrums are presented in Fig. 3 (a). The strong peak near $0.37\pi$ in the eyes closed state spectrum is due to the increased alpha rhythm.

In order to make the simulation more realistic the abrupt transition between the two stationary states was smoothed as described in [42]. In the smoothing procedure the time-varying AR coefficients having an abrupt transition at one point are presented as linear combinations of some shifted smooth basis functions. Here we have used shifted Gaussian functions. As well as the AR coefficients, also the prediction error variance evolution was smoothed.

A typical realization of the simulated ERD/ERS transition is presented in Fig. 3 (b). The simulated process was then estimated with Kalman smoother and RLS algorithms by using a fifth-order AR model. Estimates were calculated for 100 realizations and the statistics of the obtained results are presented in Table II. The bias of the averaged AR coefficient estimates and the standard deviations of single estimates are presented in the top five rows. A smaller bias for the Kalman smoother estimates is observed for the 2nd and 5th coefficient. These are the two AR coefficients for which the simulated transition effects the most. The estimates for the root corresponding to alpha rhythm were then calculated from the averaged AR coefficient estimates. Obtained estimates of the modulus and frequency of the root are presented in Fig. 3 (c). Biases

TABLE II: Statistics of different estimators for the simulated AR(5) processes. The bias and standard deviation of the estimated AR coefficients $\hat{a}_t^{(1)}, \ldots, \hat{a}_t^{(5)}$ and the bias of the modulus $|\cdot|$ and frequency $\angle\cdot$ of the root $\xi_t^\alpha$ corresponding to alpha rhythm for RLS and Kalman smoother (KS) estimators. The root $\xi_t^\alpha$ was obtained from averaged AR coefficients

| | RLS ($\lambda = 0.9$) | RLS ($\lambda = 0.98$) | KS ($\sigma_w^2 = 3 \cdot 10^{-4}$) |
|---|---|---|---|
| | Bias±STD | Bias±STD | Bias±STD |
| $\hat{a}_t^{(1)}$ | 0.063±0.079 | 0.036±0.061 | 0.130±0.074 |
| $\hat{a}_t^{(2)}$ | 0.116±0.108 | 0.156±0.081 | 0.097±0.089 |
| $\hat{a}_t^{(3)}$ | 0.108±0.127 | 0.039±0.077 | 0.182±0.111 |
| $\hat{a}_t^{(4)}$ | 0.083±0.163 | 0.040±0.109 | 0.110±0.168 |
| $\hat{a}_t^{(5)}$ | 0.239±0.152 | 0.302±0.122 | 0.165±0.137 |
| $\lvert\xi_t^\alpha\rvert$ | 0.073 | 0.126 | 0.030 |
| $\angle\xi_t^\alpha$ | 0.040 | 0.057 | 0.010 |

of the estimates are presented in bottom of Table II. RLS estimates were again calculated with two different forgetting factor values ($\lambda = 0.9$ and $\lambda = 0.98$) in order to demonstrate the trade-off between the tracking lag and stability of the estimates. From the RLS estimates it is observed that the tracking lag can not be entirely avoided even though quite small forgetting factor value is used. The Kalman smoother, on the other hand, can estimate the change very accurately.

### 3.2 Tracking of ERS/ERD transition dynamics

In order to demonstrate the tracking ability of the Kalman smoother on real EEG data we first compared the spectral dynamics calculated with different methods for a typical ERS sample. The selected sample is presented in the top of Fig. 4. Time-varying spectrum estimates obtained with Kalman smoother, RLS, LMS, and spectrogram are presented below the ERS sample. An ARMA(6,2) model was used with the adaptive algorithms. The parameters adjusting the adaptation of each method were selected to the best of our ability based on visual inspection. The state noise covariance coefficient of Kalman smoother was set to $\sigma_w^2 = 0.0003$ and the forgetting factor of RLS to $\lambda = 0.94$. The step size of LMS was adjusted adaptively to be smaller than the reciprocal of the input power. In the spectrogram a 1-second time-window corresponding to 1 Hz frequency resolution was used.

Time instant 0 in Fig. 4 corresponds to the stimulus occurrence instructing the subject to close the eyes. Shortly, after the eyes are closed, a substantial increase in alpha band power is observed in all spectra. The poor resolution of the classical spectrogram is evident when compared to the three parametric methods. The main drawback of RLS and LMS is the unavoidable tracking lag clearly seen in both spectra. The observed time-frequency resolution of the Kalman smoother is, on the other hand, extremely high. Even the short-term changes in alpha rhythm, e.g. the interval approximately from 4.3 to 4.7 seconds with decreased power in alpha frequency band, are detected reliably. Furthermore, the alpha rhythm seems to start at higher frequency and shifts to lower frequencies within few seconds. This is the so-called "squeak" phenomenon of alpha rhythm and can not be seen from the spectrogram for single trial due to poor resolution.

The ERS/ERD dynamics of all measured EEG samples were then estimated with the Kalman smoother approach. The results are presented in Figs. 5 and 6. Each analyzed subject had different characteristic patterns of alpha rhythms. Typical ERS/ERD sample and the corresponding Kalman smoother spectrum for each subject is presented in Fig. 5. For subject 1 the power of alpha rhythm is quite non-attenuating for the whole eyes closed period while for subject 3 the alpha power attenuates within few seconds after eye closure. For subject 2, on the other hand, a slowly starting increase of alpha power seems to be characteristic. The frequency band limits for alpha rhythm in Fig. 5 were adjusted based on the individual center frequencies. The center frequency

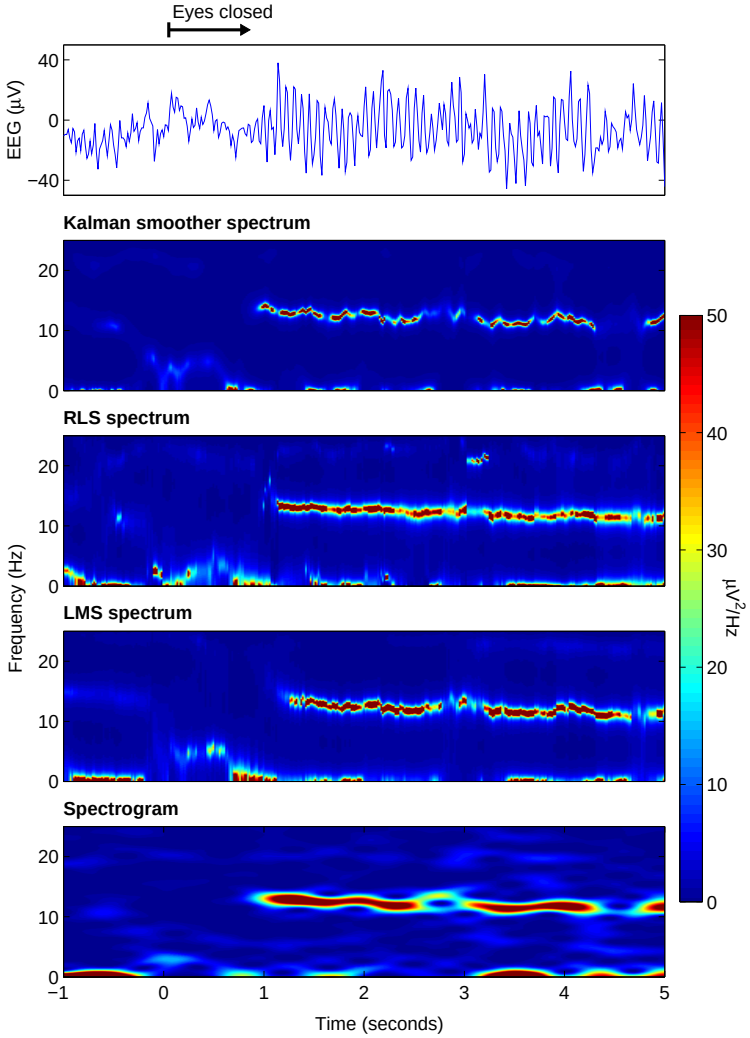


**Figure 4**: Time-varying spectral estimation of event-related synchronization dynamics of EEG alpha rhythm. The measured EEG from channel O2 is presented on the top of the figure and spectrum estimates with Kalman smoother, RLS, LMS, and spectrogram methods underneath it.

was determined as the mean peak frequency of the spectrum during eye closure. Obtained center frequencies were 12.06, 11.69, and 10.69 Hz for subjects 1, 2, and 3 respectively. The band width was set to 8 Hz.

The average ERS/ERD dynamics for each subject was then evaluated by extracting the time-variation of the center frequency and power of alpha rhythm from the Kalman smoother spectra. The center frequency was defined as the frequency bisecting the power in alpha band. The power within alpha band was obtained by simply integrating the spectrum over the alpha band. Average center frequencies and band powers over the 19 consecutive ERS/ERD samples for each subject are presented in Fig. 6 along with estimated trends and 50 % confidence intervals. The trend was estimated with a smoothness priors based approach presented in [46].

A common trend in the center frequency of alpha rhythm is observed among the subjects. Shortly after the eye closure center frequency reaches its maximum and then starts to decrease. After a while the frequency seems to reach a constant level which can be thought of as the individual characteristic alpha rhythm frequency. The power of alpha rhythm, on the other hand, increases

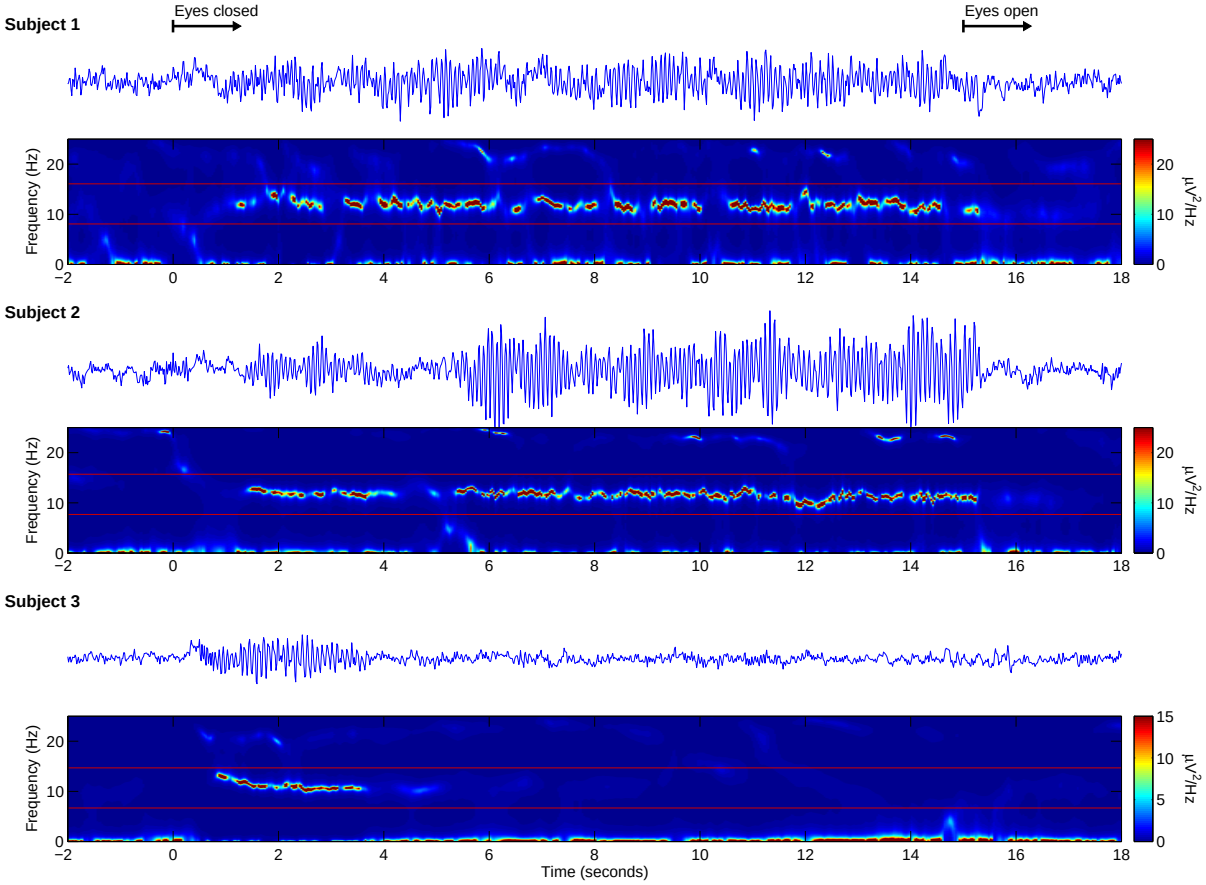


**Figure 5**: Typical ERS/ERD samples for the three selected test subjects and the corresponding time-varying Kalman smoother spectra. Stimulus occurrence times (0 and 15 seconds) instructing subjects to close and open their eyes are marked on the top. Individually adjusted alpha frequency bands are marked with horizontal lines on top of each spectrum. The band width is set to 8 Hz.

rapidly after eye closure but reaches its maximum a bit later when the center frequency has already started to decrease. After 15 seconds when the eyes are opened a rapid decrease in alpha band power is observed for subjects 1 and 2. For subject 1 the power seems to start decreasing before the eyes are opened which may be due to expectation or anticipation of the upcoming stimuli [47]. For subject 3, on the other hand, the alpha rhythm power has already been attenuated during eyes closed period.

Finally, the Kalman smoother spectrum estimation method was tested with ERS samples showing short period changes in the alpha rhythm. The results for one such sample are shown in Fig. 7. In order to prove the capability of Kalman smoother method to detect short power changes, a simple alpha detection procedure was applied. First of all, the variation of power within alpha band was calculated by integrating the Kalman smoother spectrum over the alpha band (7–13 Hz). The obtained power variation is shown in Fig. 7 (a). Then, the original EEG signal was divided into epochs for which the alpha band power was greater/smaller than the selected threshold of 5 $\mu V^2/Hz$. At last, the epochs for which alpha power was greater than the threshold and the epochs for which it was smaller were concatenated and traditional FFT based spectrum estimates were calculated for the resulting signals [Fig. 7 (b)]. The results verify the absence of alpha rhythm in the other concatenated signal and, therefore, shows the applicability of Kalman smoother spectrum estimation method in case of short period EEG changes.

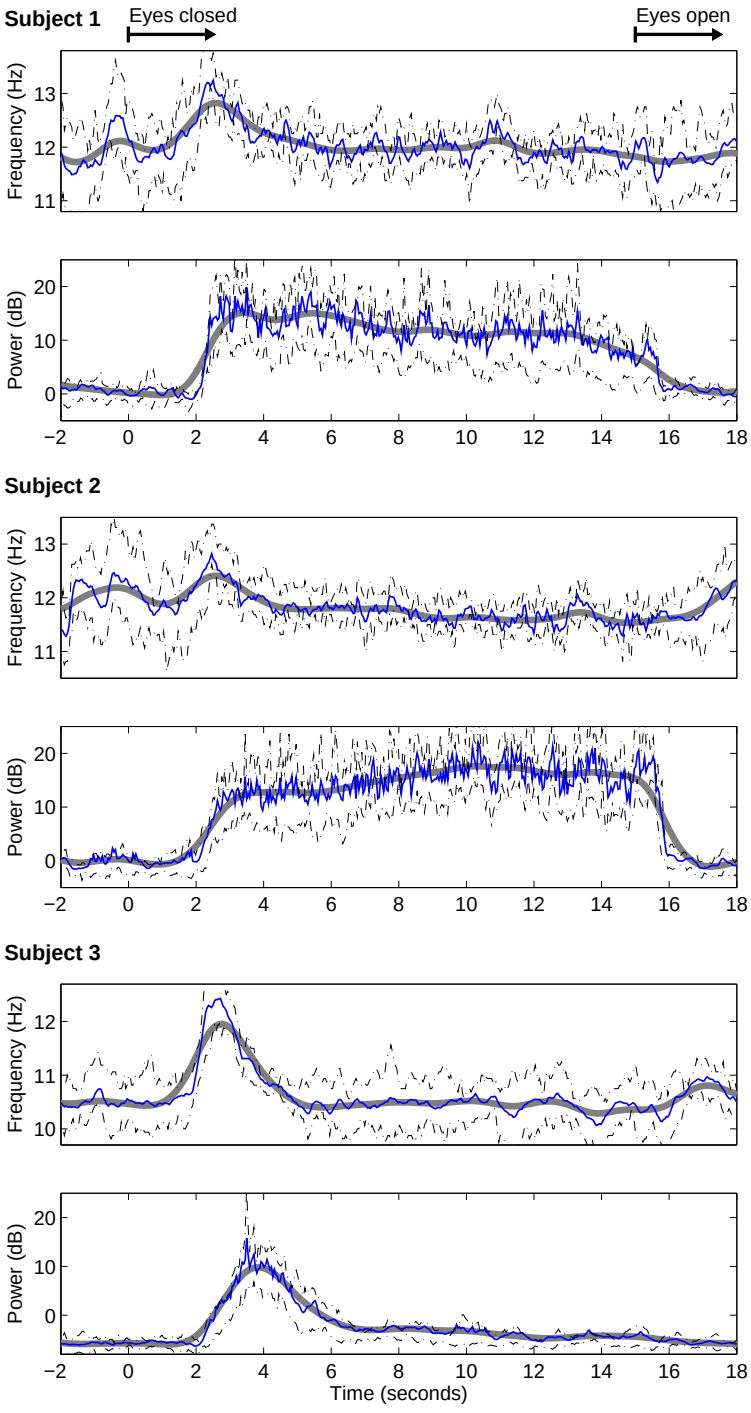


**Figure 6**: Average variation of the center frequency and band power of the alpha band for the three selected test subjects. Median values over 19 ERS/ERD samples (—), estimated trends (▬), and 50 % confidence intervals (- · -) are presented. Stimulus occurrence times (0 and 15 seconds) instructing subjects to close and open their eyes are marked on the top.

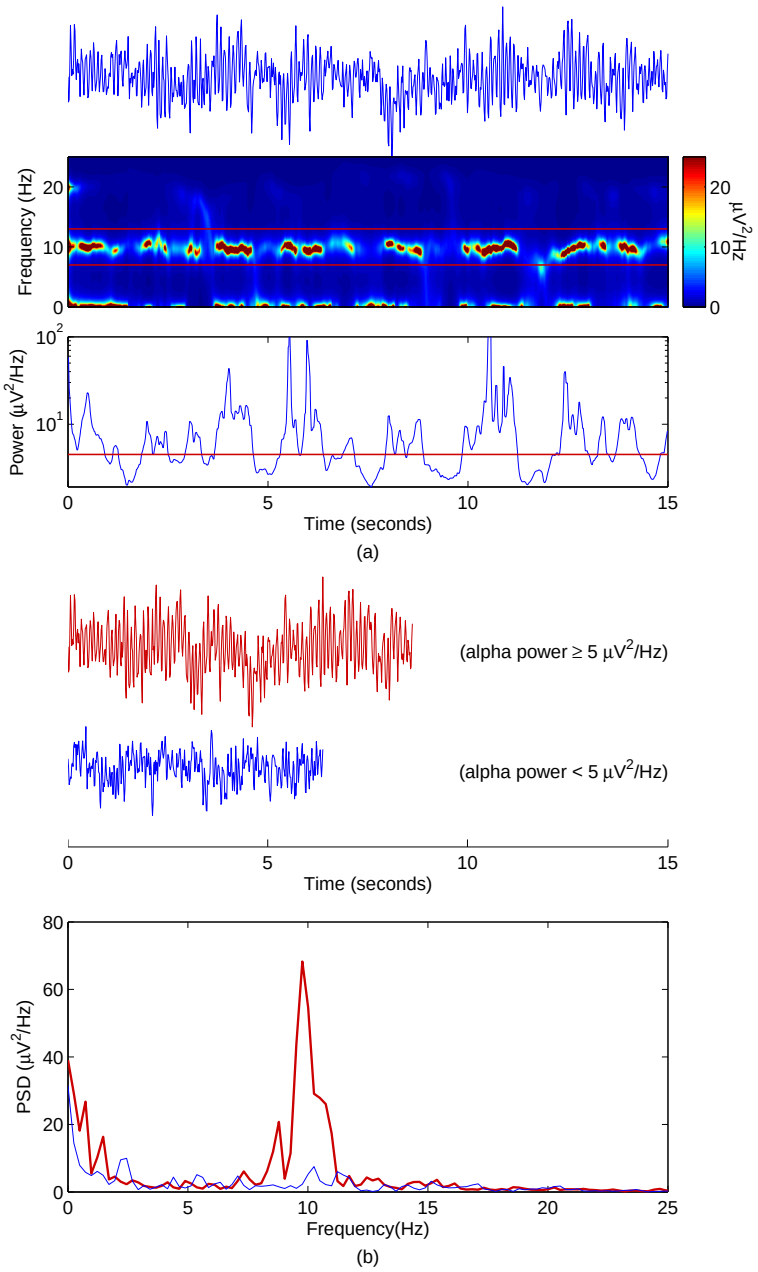


**Figure 7**: The performance of the Kalman smoother spectrum estimation method on EEG data with short period changes in alpha rhythm. (a) The EEG data measured in eyes closed state, corresponding Kalman smoother spectrum, and the power variation within the alpha band. (b) FFT based PSD estimates for the signals obtained by concatenating the EEG epochs where the detected alpha band power was greater (—) or smaller (—) than 5 $\mu V^2/Hz$.

## 4 Discussion

We have applied the Kalman filtering algorithm along with a fixed-interval smoother to tracking of nonstationary EEG. As already mentioned, the Kalman filter has been used in EEG analysis in e.g. [26, 27, 28, 29]. Kalman smoother, on the other hand, has not been previously used for analysis of EEG. Compared to other adaptive algorithms, such as RLS or LMS, the most significant benefit of the Kalman smoother is the avoidance of the time-lag in the estimates. In addition, due to the properties of parametric spectral estimation methods, the frequency resolution of Kalman smoother is better than of the classical FFT-based methods.

In this paper an autoregressive moving average model was used as a model for nonstationary EEG. ARMA model has more degrees of freedom than AR model and has, therefore, better ability to generate diverse spectral shapes. The shape of the spectrum also depends on the selected model orders. Too low a model order may result in a too smooth spectrum and, on the other hand, too high a model order may produce spurious components in the spectrum. Several model order selection criteria have been proposed for the stationary case [23], but they can hardly be applied to the time-varying case. However, based on our experience, the ARMA model of order (6,2) is suitable for modeling nonstationary EEG.

One problem in terms of application is how to set the state noise covariance coefficient adjusting the adaptation speed of the Kalman smoother. In the selection of this coefficient a trade-off between the adaptation speed and the estimate variance need to be done. In other words, the improvement in the tracking properties is achieved with the expense of the increased estimate variance. In this paper we have used a value $\sigma_w^2 = 0.0003$ for the update coefficient. The selection was made based on the conducted realistic EEG simulation and on the visual inspection of the estimates for real EEG data. Generally, the initialization of adaptive algorithms is a widely studied fundamental problem. In the case of EEG, probably the first attempt for an approach to find an optimal update coefficient for Kalman filter was made in [27]. However, there is no universal solution for this problem. The measurement noise covariance coefficient, on the other hand, was set to $\sigma_e^2 = 1$. This selection is not, however, essential since only the ratio $\sigma_e^2/\sigma_w^2$ has effect on the estimates. Furthermore, the inital guesses for the state $\theta_0$ and error covariance $C_{\tilde{\theta}_0}$ only affect on the convergence time of the algorithm and, therefore, are not essential when sufficient amount of data before the point of interest is available. Here, $\theta_0$ was set to zero and $C_{\tilde{\theta}_0}$ equal to identity matrix.

The results of tracking the ERS/ERD dynamics for the three test subjects were presented in Figs. 5 and 6. In addition, the applicability of the Kalman smoother spectrum estimation method in case of short period power changes was shown in Fig. 7. The main benefit of the presented method over the traditional ERS/ERD quantification based on bandpass filtering is the improved frequency resolution. Using the Kalman smoother approach the "squeak" effect of alpha rhythm seen in Fig. 6 for averaged estimates can be observed even from single sample spectrum estimates (see Figs. 4 and 5). With the classical spectrogram, on the other hand, this effect can not be observed from a single sample due to the poor time-frequency resolution. However, by averaging spectrograms of several consecutive ERS/ERD samples the "squeak" can be observed. The short-term "squeak" effect can not be observed from the RLS or LMS spectrums either because of the unavoidable tracking lag.

Although "squeak" effect *per se* has been suggested as clinically meaningless [48], the accuracy and usability of the Kalman smoother in the spectral analysis of event-related EEG changes makes it a promising analysis method for e.g. cognitive ERD/ERS experiments. The high time-frequency resolution of the Kalman smoother could enable the spectral analysis of single ERD/ERS samples. This property is crucial in some cognitive tasks were the transition can not be assumed to recur alike from time to time and, therefore, averaging of all trials would not be an optimal analysis method.

## References

[1] B. C. Cohen and A. Sances Jr, "Stationarity of the human electroencephalogram," *Med Biol Eng Comput*, vol. 15, pp. 513–518, 1977.

[2] L. Cohen, "Time-frequency distributions - a review," *Proc IEEE*, vol. 77, pp. 941–981, July 1989.

[3] F. Hlawatsch and G. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations," *IEEE Signal Processing Mag*, pp. 21–67, April 1992.

[4] W. Williams, "Recent advances in time-frequency representations: some theoretical foundations," in *Time Frequency and Wavelets in Biomedical Signal Processing* (M. Akay, ed.), pp. 3–43, IEEE Press, 1998.

[5] P. Durka, D. Ircha, C. Neuper, and G. Pfurtscheller, "Time-frequency microstructure of event-related EEG desynchronization and synchronization," *Med Biol Eng Comput*, vol. 39, pp. 315–321, May 2001.

[6] G. Florian and G. Pfurtscheller, "Dynamic spectral analysis of event-related EEG data," *Electroencephalogr Clin Neurophysiol*, pp. 393–396, 1995.

[7] N. Kawabata, "A nonstationary analysis of the electroencephalogram," *IEEE Trans Biomed Eng*, vol. 20, pp. 444–452, 1973.

[8] B. Graimann, J. Huggins, S. Levine, and G. Pfurtscheller, "Visualization of significant ERD/ERS patters in multichannel EEG and ECoG data," *Clin Neurophysiol*, vol. 113, pp. 43–47, 2002.

[9] S. Makeig, M. Westerfield, T. Jung, S. Enghoff, J. Townsend, E. Courchesne, and T. Sejnowski, "Dynamic brain sources of visual evoked responses," *Science*, vol. 295, pp. 690–694, January 2000.

[10] G. Pfurtscheller and F. Lopes da Silva, "Event-related EEG/MEG synchronization and desynchronization: basic principles," *Clin Neurophysiol*, vol. 110, pp. 1842–1857, 1999.

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

[12] J. Hiltunen, "Estimation of the dynamics of event-related desynchronization." Phil. Lic. thesis, University of Kuopio, Department of Applied Physics, August 1999.

[13] N. Morgan and A. Gevins, "Wigner distributions of human event-related brain potentials," *IEEE Trans Biomed Eng*, vol. 33, pp. 66–70, January 1986.

[14] N. Thakor, G. Xin-rong, S. Yi-Chun, and D. F. Hanley, "Multiresolution wavelet analysis of evoked potentials," *IEEE Trans Biomed Eng*, vol. 40, no. 11, pp. 1085–1094, 1993.

[15] I. Gath, C. Feuerstein, D. Pham, and G. Rondouin, "On the tracking of rapid dynamic changes in seizure EEG," *IEEE Trans Biomed Eng*, vol. 39, pp. 952–958, 1992.

[16] H. Zaveri, W. Williams, L. Iasemidis, and J. Sackellares, "Time-frequency representation of electrocorticograms in temporal lobe epilepsy," *IEEE Trans Biomed Eng*, vol. 39, pp. 502–509, May 1992.

[17] W. Williams, "Biological applications and interpretations of time-frequency signal analysis," in *Time Frequency and Wavelets in Biomedical Signal Processing* (M. Akay, ed.), pp. 45–72, IEEE Press, 1998.

[18] S. Schiff, J. Milton, J. Heller, and S. Weinstein, "Wavelet transforms and surrogate data for electroencephalographic spike and seizure localization," *Opt Eng*, vol. 33, no. 7, pp. 2162–2169, 1994.

[19] P. Celka, B. Boashash, and P. Colditz, "Preprosessing and time-frequency analysis of newborn EEG seizures," *IEEE Eng Med Biol Mag*, vol. 20, pp. 30–39, September/October 2001.

[20] B. Boashash and M. Mesbah, "A time-frequency approach for newborn seizure detection," *IEEE Eng Med Biol Mag*, vol. 20, pp. 54–64, September/October 2001.

[21] O. Rioul and M. Vetterli, "Wavelets and signal processing," *IEEE Signal Processing Mag*, pp. 14–38, October 1991.

[22] M. Unser and A. Aldroubi, "A review of wavelets in biomedical applications," *Proc IEEE*, vol. 84, pp. 626–638, April 1996.

[23] S. Marple, *Digital Spectral Analysis with Applications*. Prentice-Hall, 1987.

[24] R. Kalman, "A new approach to linear filtering and prediction problems," *Trans ASME J Basic Eng*, vol. 82, pp. 35–45, 1960.

[25] R. Kalman and R. Bucy, "New results in linear filtering and prediction theory," *Trans ASME J Basic Eng*, vol. 83, pp. 95–108, 1961.

[26] A. Isaksson and A. Wennberg, "Spectral properties of nonstationary EEG signals, evaluated by means of Kalman filtering: Application examples from a vigilance test," in *Quantitative Analytic Studies in Epilepsy* (P. Kellaway and I. Petersen, eds.), pp. 389–402, Raven Press, 1976.

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

[28] A. Isaksson, A. Wennberg, and L. Zetterberg, "Computer analysis of EEG signals with parametric models," *Proc IEEE*, vol. 69, pp. 451–461, 1981.

[29] M. Arnold, H. Miltner, H. Witte, R. Bauer, and C. Braun, "Adaptive AR modeling of nonstationary time series by means of Kalman filtering," *IEEE Trans Biomed Eng*, vol. 45, pp. 553–562, May 1998.

[30] J. Kaipio, *Simulation and Estimation of Nonstationary EEG*. PhD thesis, University of Kuopio, 1996. Available: http://venda.uku.fi/research/IP/.

[31] A. Schlögl, *The electroencephalogram and the adaptive autoregressive model: theory and applications*. PhD thesis, Technischen Universität Graz, 2000.

[32] A. Gelb, ed., *Applied Optimal Estimation*. The M.I.T. Press, 1974.

[33] B. Anderson and J. Moore, *Optimal Filtering*. Prentice Hall, 1979.

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

[35] S. Haykin, *Adaptive Filter Theory*. Englewood Cliff, NJ: Prentice-Hall Inc., 1986.

[36] J. Melsa and D. Cohn, *Decision and Estimation Theory*. McGraw-Hill, 1978.

[37] P. Karjalainen, "Estimation theoretical background of root tracking algorithms with application to EEG." Phil. Lic. thesis, University of Kuopio, Department of Applied Physics, 1996. Available: http://venda.uku.fi/research/biosignal/publications/.

[38] P. Karjalainen, *Regularization and Bayesian methods for evoked potential esimation*. PhD thesis, University of Kuopio, Department of Applied Physics, 1997. Available: http://venda.uku.fi/research/biosignal/publications/.

[39] L. Ljung and S. Gunnarsson, "Adaptation and tracking in system identification – a survey," *Automatica*, vol. 26, pp. 7–21, 1990.

[40] L. Ljung, "General structure of adaptive algorithms: adaptation and tracking," Tech. Rep. LiTH-ISY-I-1294, Linköping University, December 1991.

[41] A. Sayed and T. Kailath, "A state-space approach to adaptive RLS filtering," *IEEE Signal Processing Magazine*, vol. 11, pp. 18–60, 1994.

[42] J. Kaipio and P. Karjalainen, "Simulation of nonstationary EEG," *Biol Cybern*, vol. 76, pp. 349–356, 1997.

[43] A. Isaksson and A. Wennberg, "An EEG simulator - a means of objective clinical interpretation of EEG," *Electroencephalogr Clin Neurophysiol*, vol. 39, pp. 313–320, 1975.

[44] A. Wennberg and A. Isaksson, "Simulation of nonstationary EEG signals as a means of objective clinical interpretation of EEG," in *Quantitative Analytical Studies in Epilepsy*, pp. 493–509, Raven Press, 1976.

[45] M. Nuwer, "Quantitative EEG: I. techniques and problems of frequency analysis and topographic mapping," *J Clin Neurophysiol*, vol. 5, no. 1, pp. 1–43, 1988.

[46] M. Tarvainen, P. Ranta-aho, and P. Karjalainen, "An advanced detrending method with application to HRV analysis," *IEEE Trans Biomed Eng*, vol. 49, pp. 172–175, February 2001.

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

[48] B. Westmoreland and D. Klass, "Unusual EEG patterns," *J Clin Neurophysiol*, vol. 7, no. 2, pp. 209–228, 1990.