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

Simulation of Nonstationary EEG

J.P. Kaipio P.A. Karjalainen

June 14, 1994

Report No. 6/94

This manuscript has been submitted to *Biological Cybernetics* on June 17, 1994

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

Simulation of Nonstationary EEG

J.P. Kaipio P.A. Karjalainen

June 14, 1994

Abstract

In this paper we present a systematic method to generate realizations of nonstationary EEG. The method is based on modeling EEG first as a state evolution. The states are initially represented as segments of stationary AR processes which are described with the corresponding predictor coefficients and prediction error variances. These parameters are then concatenated to give a piecewise time-invariant parameter evolution. The evolution is projected onto an appropriately selected set of smoothly time-varying functions. This projection is used to generate the realization. As an example we use this method to simulate the EEG of a drowsy rat. This EEG can be described as toggling between two states that differ in the degree of synchronization of the activity inducing neuron clusters.

1 Introduction

Since the operations of estimation and simulation of time series are inverse to each other, we will discuss them in this paper in parallel. The end to the simulation is thought to be the estimation of the time-varying characteristics of the process and to offer meaningful references to the estimates.

1.1 The estimation and simulation of nonstationary processes

In a real world setting we are hardly ever able to assess the performance of a method by just applying it to observed data (a sample). In many cases a method to be tested is compared with an old one that is assumed to give the correct answers and thus serves as a reference. In the estimation of properties of stationary processes we have several methods with which we can obtain consistent estimates if the process is ergodic. In this case the time averages can be used to estimate the ensemble averages [18]. We can then make our estimates more accurate by using more data to calculate them. The accurate estimates can subsequently serve as references against which the methods to be tested a compared.

In the case of nonstationary processes we often cannot use time averages. To estimate a property of the process at a time instant we have, in the strict sense, just one point of the sample to build our estimate on. To get over this problem the assumption of slow time variation is often adopted. This gives us a possibility to use *e.g.* adaptive segmentation which will chop the sample to quasi-stationary (almost stationary) epochs to which we would then apply methods that are intended to be used with samples from ergodic processes [3, 11]. The applicability of this approach depends heavily on the slowness of time variation, the criteria on which the segmentation is based and the correctness, or rather, the adequacy of the time series model.

There are cases in which the time variation is so rapid that the process cannot be approximated with a concatenation of stationary epochs. The lengths of the quasi-stationary segments would then be too short to enable adequate estimates. This applies to both estimation and simulation of such processes. On the other hand, we cannot allow for a time variation that is totally "unpredictable". This would mean that if we have obtained a sample of the process and aim to estimate its time-varying characteristics, the model may have, loosely speaking, more degrees of freedom than the sample the estimates are based on [9].

If we can make an assumption of *smoothness* of the time variation we can still obtain meaningful estimates and at the same time allow reasonably rapid changes of process characteristics. This follows from the fact that one can have a smooth sequence of parameters describing the process,

i.e. a sequence with small second differences, that has large first differences. This means that the parameter rate of change does not have to be small as such. This applies also to the building of models for time-varying processes and the use of these models to the generation of realizations of these processes.

1.2 Autoregressive models of EEG

The autoregressive (AR) process model was first applied to the EEG in 1969 by Zetterberg [23], Fenwick *et. al* [5] and Gersch [7]. Since that time the AR model has been applied to the EEG in at least 160 journal articles [12]. The applicability of AR models to the EEG has usually been verified only experimentally but recently there has been some interest towards the theoretical justification of this [22, 2]. AR models have been used in a wide collection of tasks concerning EEG, including spectral analysis, classification of brain states, detection of transients, prediction of epileptic seizures and the simulation of EEG.

The simulation of EEG using an AR model was suggested early in the 70' by several researchers [4, 21, 24]. Simulations of time-varying AR processes were used by Wennberg and Isaksson [20] to test the modeling and tracking capabilities of the Kalman filter. The aim in this paper was to assess the applicability of the Kalman filter to the estimation of the time-varying EEG and the simulations were made to approximate the classical band structure of EEG. They used an approximate factorization of AR model to first and second order spectral factors [23] which have been shown to correspond to the classical bands of EEG. To produce time-varying realizations they let the parameters of these factors, *i.e.* the power, bandwidth and the center frequency, to be time-varying. They did not, however, discuss how the time variability should be systematically implemented to produce realizations that would exhibit the time-varying characteristics of the EEG. In addition, the classical band structure of EEG does not always apply, especially when the EEG exhibits at least partly chaotic characteristics. They also discussed the use of simulations to aid in the recognisation of certain spectral parameters of the model using visual inspection [10].

In this paper we suggest a systematic method which can be used to generate realizations for the simulation of time-varying EEG. The method is based on dividing the EEG to classes that are described by estimated distributions of their predictor (AR model) coefficients and prediction error variances. If the class evolution is not predetermined, this evolution is then simulated using *e.g.* estimated state transition probabilities and Markov models or state lengths (durations). The parameter representers of the classes are then drawn from the respective distributions according to the state evolution and concatenated in time. This piecewise stationary parameter evolution is projected to a set of smoothly time-varying functions, which projection is finally used to generate the realizations by time-varying filtering.

2 Estimation and simulation of the state evolution and class representers

We assume that the time-varying EEG can be approximated with a progression of states $s_j,\ j = 1,\ldots$ each of which is a member of a class $\mathcal{C}_k,\ k = 1,\ldots,K$, and that within these states the process can be approximated with an AR(p) process. We also assume that within each state the coefficients $\phi_i^j,\ i = 1,\ldots,p$ of the predictor and the prediction error variance γ_j corresponding to the AR(p) model are drawn from the density that can be approximated with a $p+1$ –dimensional normal distribution. Each class has its own distribution. We call the vector $\theta^j = (\phi_1^j,\ldots,\phi_p^j,\gamma_j)'$ a *representer* of the state s_j. We will further assume that the representer of the state s_j is independent of the representers of the states s_l for all $j \neq l$. We will use the term state evolution also when it could be more correct to refer to class evolution.

There are several methods with which to estimate such a state evolution if the (probabilistic) description is not known *a priori* . One such a method is to model the state evolution as a Markov model and estimate the parameters of the model. This model can then used this model to simulate the state evolution. The estimation of a continuous-time Markov model and the simulation of sleep EEG state evolution has been discussed in [13].

The assumption of approximate normality of representers implies that we only need to estimate the means and covariances of θ^k for each class $\mathcal{C}_k$ to be able to generate representers.

If the means and covariances of θ^k are not known, they can be estimated from segmented and classified EEG data *e.g.* by using the modified covariance method to calculate the parameters of the AR(p) model and the associated residual error variance [17]. These can be used *in lieu* of the predictor coefficients and the prediction error variance.

Let there be $n_k > p$ estimates $\hat{\theta}^k$ for θ^k. Estimates for the mean $\hat{\mu}_k$ and covariance $\hat{C}_k$ of $\hat{\theta}^k$ can then be obtained as

$$\hat{\mu}_k = n_k^{-1} \sum_{i=1}^{n_k} \hat{\theta}_i^k \tag{1}$$

$$\hat{C}_k = (n_k - 1)^{-1} \sum_{i=1}^{n_k} (\hat{\theta}_i^k - \hat{\mu}_k)(\hat{\theta}_i^k - \hat{\mu}_k)' \tag{2}$$

We can now generate representers from the thus estimated $p+1$-variate normal distributions

$$f_k(\theta) = (2\pi)^{-p/2} |\hat{C}_k|^{-p/2} \exp\left(-(\theta - \hat{\mu}_k)' \hat{C}_k^{-1} (\theta - \hat{\mu}_k)\right) \tag{3}$$

as follows: Generate a $p+1$-vector g of independent normal variables with zero mean and unit variance. Then the variables $\theta = L_k g + \mu_k$ are distributed as $f_k(\theta)$, where $\hat{C}_k = L_k L_k'$ is the Cholesky decomposition of the $\hat{C}_k$ [18].

Note that if the representers have been estimated from relatively short segments, the covariances $\hat{C}_k$ may include a non-negligible contribution that is due to the small sample estimate properties. In such a case we can reduce the variances of the classes simply by dividing the covariance matrices by a number that is greater than unity. While this method is not the correct way to diminish the small sample contribution, it maintains the eigenvector structure of $\hat{C}_k$ and only diminishes the eigenvalues. This means that the only result of this operation is that the deviances of the simulated θ^ks from the means $\hat{\mu}_k$ will be smaller. This will also decrease the possibility of obtaining a temporarily unstable model.

An AR(p) model is stable if and only if the roots $z_i,\ i = 1, \ldots, p$ of the polynomial

$$X(z, \phi) = 1 - \sum_{i=1}^{p} \phi_i z^{-i} \tag{4}$$

fulfil $|z_i| < 1$ for all t. For some classes the roots of $X(z, \phi)$ with $\phi_i = \mu_i,\ i = 1, \ldots, p$ may have almost unit modulus or the covariance may be large, which means that the probability of obtaining an unstable representer is non-negligible. This is why the stability of each representer should be verified by calculating the roots of the polynomial. All roots with modulus greater than or equal to unity should be transferred to lie inside the unit circle and the predictor reassembled. This can be done by factoring the polynomial to (complex) roots and multiplying each unstable root z_i by $|z_i|^{-2}$. This operation will approximately restore the shape of the spectrum in most practical cases [14, subroutine fstab].

3 Generation of the time-varying AR model

We will discuss first two types of direct concatenation that are not feasible to generate time-varying realizations and how this unfeasibility can be overcome. Then we discuss the details of the method that was shortly described in the end of Section 1.

3.1 Direct concatenation

The majority of time series methods that are designed to track the characteristics of time-varying processes are based on either direct or indirect minimization of the prediction (residual) error [15]. It is thus desirable that the optimal predictor of the realization behaves in a way that it makes sense to compare the estimated model with it. For example, if the coefficients $b_k(t)$ of the optimal

predictor will decay relatively fast, we would use as the definition of time-varying spectral density [9]

$$P(\omega,t)=\frac{\sigma_e^2(t)}{\left|1-\sum_{k=1}^{\infty}b_k(t)\exp(-i\omega kT)\right|^2} \tag{5}$$

There are two trivial ways to use concatenation to generate realizations of nonstationary processes which ways are not directly usable. The first is to generate separately stationary segments and then concatenate them directly. There are two major problems with this kind of realizations. In this case we have actually a non-overlapping sum of independent processes. Let $x(t)$ be the realization with zero mean for all t and t^* be the first instant of a new segment. Then, irrespective of the values at time t^*-1, the correlation of $x(t^*)$ with the entire history of $x(t)$ is zero, the predictor for $x(t^*)$ is zero and the spectrum estimate $P(\omega,t^*)=\sigma_e^2(t^*)$. This is clearly not desirable.

On the neuronal level this kind of a sudden change in process characteristics would correspond to a situation where one activity inducing neuron cluster would shut down abruptly and at the same time another cluster would gain its full effect. This is a totally unrealistic situation.

The second type of concatenation involves the concatenation of the parametric representations of the segments. This concatenation could then be used as a time-varying filter with abrupt filter coefficient changes at the segment borders. As above, this kind of change would correspond to an unrealistic abrupt change in the physical and chemical state of the neuron cluster. In addition, it turns out that the calculation of the optimal predictor is far from straightforward.

For this reason we have chosen to build a smoothly changing parametric representation for the process. We use a time-varying AR(p) model as the representation and generate the realizations by feeding white noise to the corresponding filter. If the coefficients of the model do not change very much during the correlation time of the process, the optimal predictor is approximately equal to the coefficients of the AR(p) at each time. Thus the reference against which the estimates are compared is directly accessible.

3.2 The generation of smooth parameter evolution

We start from a sequence of representers which are described as AR(p) models and represented by the corresponding predictor coefficients and prediction error variances $\phi_k^i,\ k=1,\ldots,p$ and γ_i, respectively. The corresponding sequence of the lengths of the segments is $T_i,\ i=1,\ldots$. We call the concatenate of the segments a block which is of length $T=\sum T_i$. Build the coefficient matrix Φ^{c} of piecewise constant parameter evolution

$$\Phi^{\mathrm{c}}=\begin{pmatrix}
\left.\begin{matrix}\phi_1^1 & \phi_2^1 & \cdots & \phi_p^1\\ \vdots & & & \vdots\end{matrix}\right\}T_1\\
\left.\begin{matrix}\phi_1^1 & \phi_2^1 & \cdots & \phi_p^1\\ \phi_1^2 & \phi_2^2 & \cdots & \phi_p^2\\ \vdots & & & \vdots\end{matrix}\right\}T_2\\
\begin{matrix}\phi_1^2 & \phi_2^2 & \cdots & \phi_p^2\\ \phi_1^3 & \phi_2^3 & \cdots & \phi_p^3\\ \vdots & & & \vdots\end{matrix}
\end{pmatrix}=[\Phi_1^{\mathrm{c}}\ \Phi_2^{\mathrm{c}}\ \ldots] \tag{6}$$

Next we smooth each individual parameter process $\Phi_j^{\mathrm{c}}\ j=1,\ldots,p$ by projecting them separately onto a smooth subspace $\mathcal{S}_{\mathrm{s}}$ of $\mathbb{R}^T$, where $T=\sum T_i$. Let us for the moment suppose, that the smooth subspace $\mathcal{S}_{\mathrm{s}}$ is such that there are projections onto it which restore the approximate "shapes" of Φ_j^{c}. Let now $S=[\varphi_1,\varphi_2,\ldots,\varphi_M]$ be a $T\times M$ matrix containing a basis of $\mathcal{S}_{\mathrm{s}}$ in its columns.

The orthogonal projections Φ_j of Φ_j^{c} onto $\mathcal{S}_{\mathrm{s}}$ are

$$\Phi_j=\left.\Phi_j^{\mathrm{c}}\right|_{\mathcal{S}_{\mathrm{s}}}=S(S'S)^{-1}S'\Phi_j^{\mathrm{c}} \tag{7}$$

The invertibility of $S'S$ is trivial since S is assumed to contain a basis for $\mathcal{S}_\mathrm{s}$. The projection can be performed simultaneously for all j giving

$$\Phi = S(S'S)^{-1}S'\Phi^\mathrm{c} \tag{8}$$

The predictor coefficients $\phi_k(t)$ are now of the form

$$\phi_k(t) = \sum_{l=1}^{M} q_{kl}\varphi_l(t) \tag{9}$$

where $\{q_{kl}\} = Q = (S'S)^{-1}S'\Phi^\mathrm{c}$. To complete the specification we need the correspondingly smoothed evolution of the prediction error variance. As above, concatenate γ_i to give

$$\Gamma^\mathrm{c} = (\overbrace{\gamma_1, \ldots, \gamma_1}^{T_1}, \overbrace{\gamma_2, \ldots, \gamma_2}^{T_2}, \gamma_3, \ldots)' \tag{10}$$

Then using the same basis we obtain

$$\Gamma = S(S'S)^{-1}S'\Gamma^\mathrm{c} \tag{11}$$

Next we generate a white noise sequence of zero mean and unit variance $e(t),\ t = 1, \ldots, T$. The realization is then obtained as

$$x(t) = \sum_{k=1}^{p} \Phi(t,k)x(t-k) + \sqrt{\Gamma(t)}e(t) \tag{12}$$

Due to the highly nonlinear mapping of polynomial coefficients to the roots, it would be very burdensome to add the stability condition $|z_k(t)| < 1$ to the projection as a constraint in the least squares problem corresponding to the projection. This constrained least squares problem can be expressed as

$$\min_{\Phi}\{\|\Phi - \Phi^\mathrm{c}\|,\ \Phi_j \in \mathcal{S}_\mathrm{s}\ \forall j,\ |z_k(t)| < 1\ \forall k, t\} \tag{13}$$

Since Φ tends to overshoot the stationary representers in parameter space, it is probable that in some cases the roots $z_k(t)$ will be temporarily unstable. This is, however, not a problem if the overshoot and its extent in time are not large since in this case $x(t)$ does not diverge much.

3.3 The choice of basis for $\mathcal{S}_\mathrm{s}$

It is always necessary to include the constant basis as one of the functions so that the stationary case falls into this setting as a special case whatever the basis. For convenience we will thus assume that $\varphi_1 \equiv 1$.

The basis function selection has been dicussed in the literature concerning time-varying AR modeling, for a review see [8]. The problem of estimating optimal time-varying AR (TVAR) model can be viewed to be the inverse to the one at hand. In TVAR modeling the parameter evolution is constrained so that each parameter process is in $\mathcal{S}_\mathrm{s}$.

The basis functions suggested to be used in conjunction with TVAR models include general polynomial, Fourier, Haar, Legendre, spline and prolate spheroidal wave bases. With the exception of the Haar basis, [1] which is a non-overlapping block pulse basis, the suggested bases are able to model smooth changes in the parameter evolution. This ability is dependent on the choice of basis and the parameters concerning it, the dimension of $\mathcal{S}_\mathrm{s}$ in particular.

The problem with these bases is that they are global within the block so that a small change anywhere in the block induces changes that extend overall within the block. This is obviously not desirable. What we wish to have is to maintain the local nature of the Haar basis and simultaneously obtain a smooth parameter evolution.

[1] The fitting of Haar basis is equivalent to segmenting the data and estimating the least squares solutions for the segments individually.

We have chosen to use a basis consisting of shifted and scaled (sampled) Gaussian functions $s_i(t)$:

$$\varphi_i(t) = \exp\left(-(t-t_i)^2/d_i^2\right), \quad i = 2, \ldots, M \tag{14}$$

Obviously there can be no fixed general choice for M, t_i or d_i that would be optimal in some sense. The choice of these parameters depends heavily on the underlying situation and the demands. In the case of simulation the feasible parameters depend on the number and lengths (durations) of the stationary segments, or more generally, the distribution of these two.

According to our experience the Gaussian basis fulfils the requirements of locality and smoothness of evolution with relatively simple selection rules of M, t_i or d_i for a wide class of evolutions. The trivial choice for t_i and d_i is such that the half-widths of φ_i will exactly cover the length of the block. This leads to $t_i = (i-1)T/(M-1)$, $2Md_i = T$, where T is the length of the block. The selection of M is yet left open. If the block consists of D segments of length L each, then $T = DL$ and we could choose $M = D + 1$. This selection leads to projections that are smooth but tend to overshoot the design classes heavily in the parameter space.

We have found that the selection

$$M \approx 2D \tag{15}$$

$$d_i \approx T/4M \tag{16}$$

$$t_i = (i-1)T(M-1) \tag{17}$$

will produce reasonable projections if the segment length distribution is not very wide.

Other methods to obtain a smooth parameter evolution

Naturally it is possible to tailor the bases individually to follow the state evolution and just to smooth the transients of the parameter evolution. One way to implement this indirectly is to use the sigmoidal basis. An example of sigmoidal type basis functions is

$$\varphi_i(t) = (1 + \exp(-c(t-t_i)))^{-1} \tag{18}$$

where t_i are selected to coincide with the segment borders and c is adjusted to give desired rise times.

It is also easy to build Φ^c and then filter the columns with a (noncausal) zero-phase low-pass filter. Just like with the sigmoidal basis this will, however, lead to a situation where the realization would be a concatenate of stationary and transition regions which is not very realistic.

In addition, it has been shown that the tracking capability of adaptive estimators depends on the type of evolution of the parameters [1]. If the parameters would be time-invariant for a long period of time, the estimators could perform markedly better with the realization than with the real EEG data.

In principle it is possible to realize the parameter evolution as a multivariable AR process. For many situtions this could be assumed to be a feasible model for parameter evolution. The probability density function of the parameters is, however, very difficult to control so that the model remains stable at all times. If the parameter processes are forced to the stability region after generation, the smoothness exhibited *e.g.* by a second order low pass type AR process is difficult to maintain. The testing for stability must be carried out for each t which will be very burdensome with long realizations.

4 An example

As an example of the use of the proposed method we simulate the EEG of a drowsy rat. This is a simple example since this EEG can be approximated as a process that toggles between two states. The amount of data on which the statistics we give here is far too small to enable any true inference. The main point of this example is to illustrate the steps of the proposed method in detail.

The state evolution

The data was visually classified to two states. As an aid in the classification the spectrum estimates and root locations of the modified covariance estimates of AR(6) model [17].

The exponential distribution induced by the Markov model did not fit the experimental segment length distribution of either class. We adopted another model for the state evolution. Since there are only two states, state 1 always changes to state 2 and *vice versa* . Thus we can model the state evolution simply by the estimated distributions of the segment lengths. The gamma distribution

$$g(T; \lambda_1, \lambda_2) = \frac{\lambda_2^{\lambda_1+1}}{G(\lambda_1 + 1)} T^{\lambda_1} \exp(-\lambda_2 T) \tag{19}$$

seemed to fit the observed histograms of both states, where G is the gamma function. The parameters of the distributions were estimated using the BFGS quasi-Newton method to find the maximum likelihood estimates. The estimated probability densities $g_1(T) = g(T; \lambda_1^1, \lambda_2^1)$ and $g_2(T) = g(T; \lambda_1^2, \lambda_2^2)$ are shown in Fig 1.

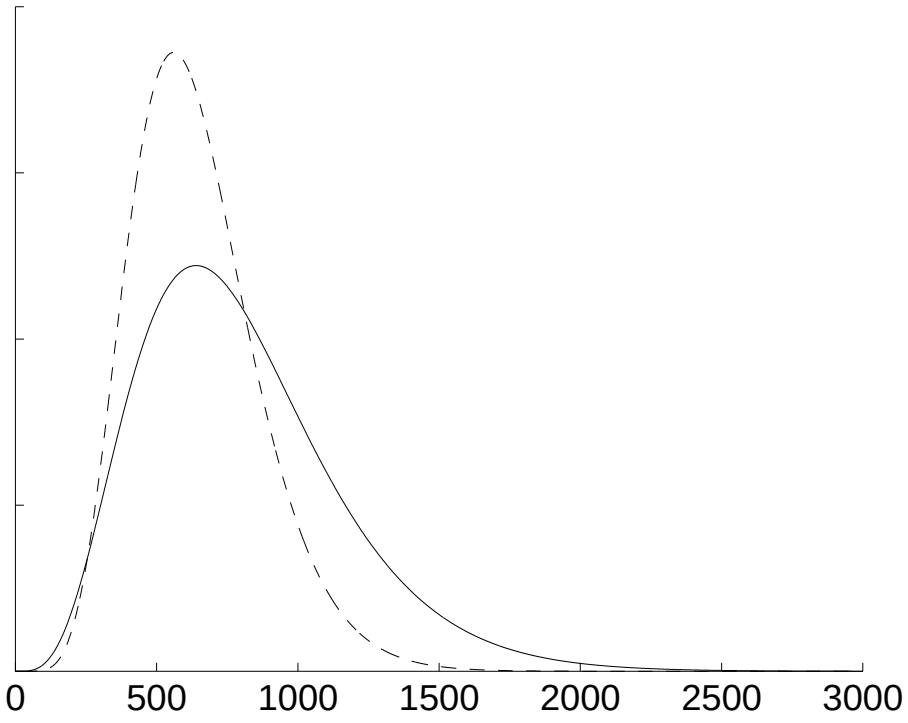

Figure 1: The maximum likelihood estimates $g_1(T) = g(T; 4, 0.006)$ and $g_2(T) = g(T; 7, 0.013)$ of the segment length densities for classes 1 (solid) and 2 (dotted).

The representer estimates

We used the forward-backward least squares method to estimate the coefficients of AR(6) models for both states. These coefficients and the corresponding residual variances were used as estimates of the predictor coefficients and prediction error variances. Forty estimates of the representers θ^1 and θ^2 were obtained. The marginal densities of each $\phi_i,\ i = 1, \ldots, 6$ were visually examined to support the normal approximation of the predictor coefficients. The marginal density of the

prediction error variances γ, however, could not be approximated with the normal distribution, which was to be expected. To use another distribution would be very cumbersome, so we let the joint density be normal but applied hard limits to γ when representers were drawn from the distribution. As hard limits we used the minimum and maximum of the observed residual variances of both classes.

The means $\hat{\mu}_1$, $\hat{\mu}_2$ and the covariances $\hat{C}_1$ and $\hat{C}_2$ are given in Table 1. The spectra and root locations corresponding to the means are given in Fig. 2. Note that the frequency corresponding to the second peak in the spectrum of state 2 is approximately two times the frequency of the first. This can be taken as a sign of partially chaotic nature of state 2.

TABLE 1
The estimated means $\hat{\mu}_1$, $\hat{\mu}_2$ and the covariances $\hat{C}_1$ and $\hat{C}_2$ of classes 1 and 2.

	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	ϕ_6	γ
$\hat{\mu}_1$	0.5065	-0.0528	0.0619	-0.0828	-0.0009	-0.1431	0.5686
$\hat{\mu}_2$	0.6722	-0.3583	-0.1741	-0.0562	-0.0344	-0.2811	1.1819
$\hat{C}_1$	0.0362	-0.0335	0.0182	-0.0016	0.0049	-0.0055	0.0223
	-0.0335	0.0448	-0.0185	-0.0033	-0.0027	0.0031	-0.0296
	0.0182	-0.0185	0.0283	-0.0029	-0.0034	-0.0005	0.0056
	-0.0016	-0.0033	-0.0029	0.0139	0.0003	-0.0063	-0.0054
	0.0049	-0.0027	-0.0034	0.0003	0.0182	-0.0124	0.0026
	-0.0055	0.0031	-0.0005	-0.0063	-0.0124	0.0183	0.0082
	0.0223	-0.0296	0.0056	-0.0054	0.0026	0.0082	0.0444
$\hat{C}_2$	0.0177	-0.0066	-0.0101	0.0127	0.0073	-0.0046	-0.0119
	-0.0066	0.0179	-0.0115	0.0111	-0.0092	0.0097	-0.0072
	-0.0101	-0.0115	0.0398	-0.0343	-0.0026	0.0157	0.0083
	0.0127	0.0111	-0.0343	0.0408	-0.0070	-0.0021	-0.0175
	0.0073	-0.0092	-0.0026	-0.0070	0.0197	-0.0192	0.0066
	-0.0046	0.0097	0.0157	-0.0021	-0.0192	0.0359	-0.0013
	-0.0119	-0.0072	0.0083	-0.0175	0.0066	-0.0013	0.4413

The generation of realizations

The segment length sequence was generated by drawing independent random numbers from the distributions g_1 and g_2 in turn. The representer sequence was generated as described in Section 2.

The segment lengths and representers of a realization are given in Table 2. The corresponding concatenated predictor coefficient processes (vectors) Φ_1^c and Φ_2^c and the corresponding smoothed processes Φ_1 and Φ_2 are shown in Fig. 3a. The prediction error variances Γ^c and Γ are shown in Fig. 3b. The corresponding realization is shown in Fig. 3c and an example of the original EEG in Fig. 3d.

To verify the similarity of the quasi-stationary epochs and the transition region between the original rat EEG data and the realizations, we show in Fig. 4a a segment of original data and a realization in Fig. 4b which is adapted to the original segment as follows: The EEG was divided visually into three parts and the representers for the quasistationary parts were estimated. The temporal extent of the representers was broadened to cover the transition region. The concatenated and smoothed parameter evolution is shown in Fig. 4c. As we can see, the stationary phases and the transition of the simulation are virtually indistinguishable.

5 Discussion

We have presented a systematic method which can be used to simulate nonstationary EEG. The applicability of the method to *e.g.* human sleep EEG simulation is obvious. Apart from EEG that

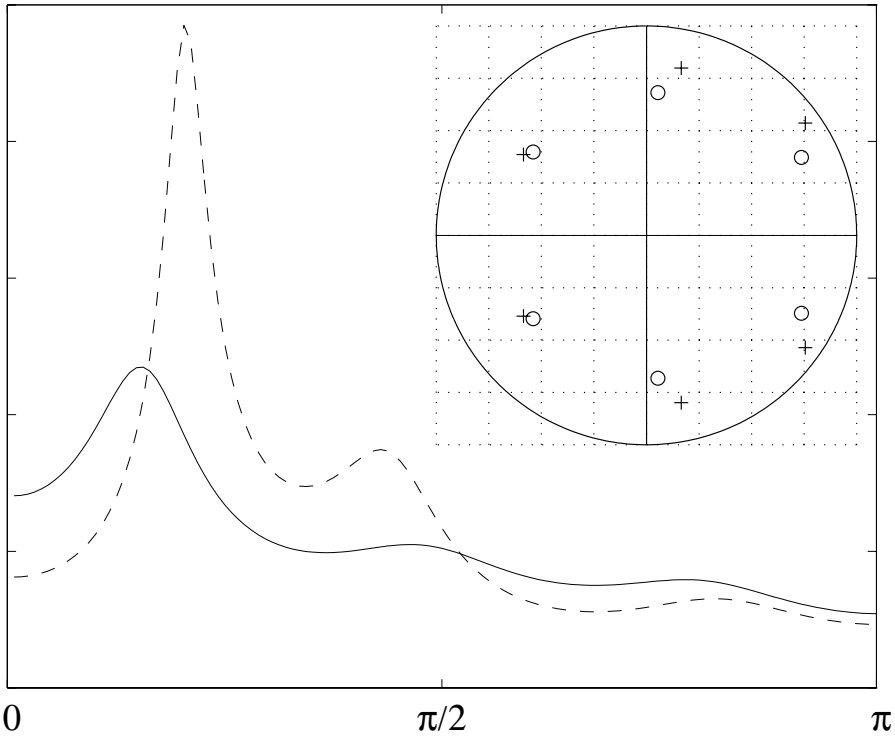

Figure 2: The spectra of classes 1 (solid) and 2 (dotted) corresponding to the means μ_1 and μ_2 and the model polynomial root (pole) locations in the complex plane, $\circ$: class 1, $+$: class 2

can be described by a stochastic state evolution, this method is also applicable to such processes that exhibit a deterministic type state evolution.

An example is the gradual desynchronization of alpha waves as a response to visual stimulus [16]. In such a case we can take samples of the EEG that are synchronized with the stimulus, estimate representer statistics before the stimulus using one segment $(k_1; T_1)$ and after the stimulus using several segments $(k_2, k_2, \ldots; T_2, T_3, \ldots)$. We can then use the predetermined state order $k_1, k_2, \ldots$ and segment lengths and vary the representers only.

Signals such as EOG, EMG and some event related potentials can be added to the realizations directly. When doing this, however, it must be noted that the occurrences may correlate some states of the background. An example is the correlation of EOG with REM sleep. To achieve a realistic situation the occurrence statistics should be estimated and used accordingly.

It is well known that some epochs of EEG are better described as a mixture of chaotic and stochastic behaviour rather than a regular stochastic process such as an AR(p) model, see *e.g.* [19]. However, for short segments the main difference between chaotic and regular description is that the former exhibits phase correlation whereas the phase of the latter should be independent between any two different frequencies and have uniform distribution between zero and 2π. Short segments of chaotic processes are periodic and can thus be modeled as a limiting case of a regular process having a line spectra with the lines exactly at the multiples of the inverse of the period. Such a process can be approximated also with an AR(p) model. As an example we can draw into attention the state 2 of the rat EEG and the corresponding simulation of Fig. 4. The visual appearance of the chaotic state 2 does not differ much from the simulation that is regular by construction. See [2] for a discussion of the applicability of AR(p) model to chaotic EEG.

If this approach is not considered adequate, the prediction error process Γ can be smoothly forced to zero and the thus obtained gap filled by a chaotic process that is multiplied with a taper

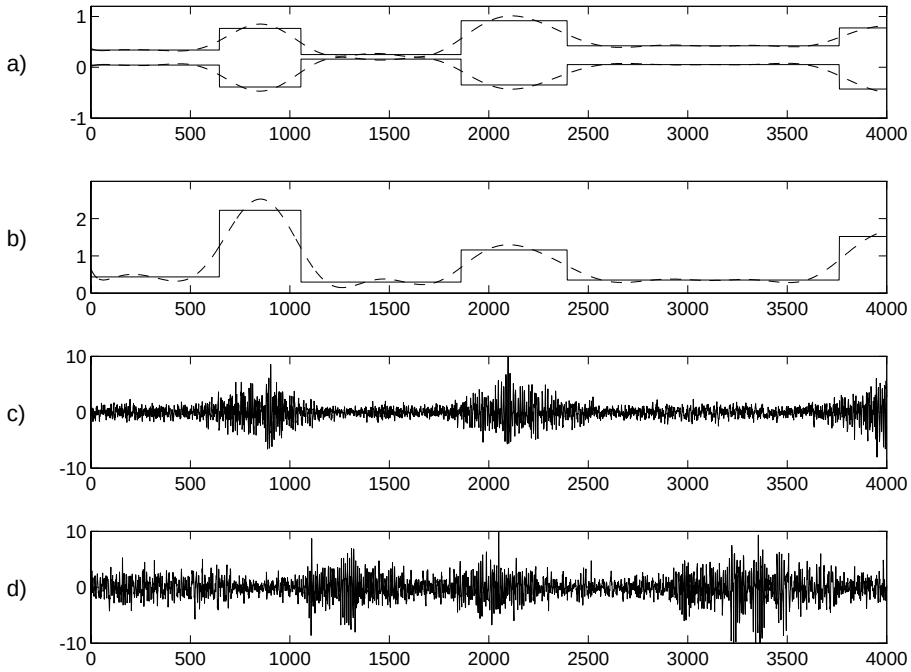

Figure 3: a) The concatenated predictor coefficient processes Φ_1^c and Φ_2^c and the corresponding smoothed processes Φ_1 and Φ_2. b) As in a) but for prediction error variances Γ^c and Γ. c) The corresponding realization. d) An example of original rat EEG.

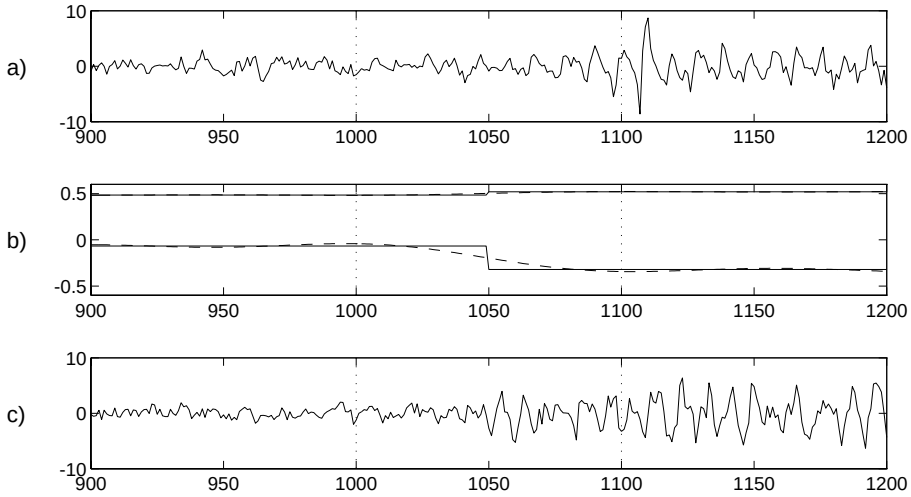

Figure 4: a) A segment of original rat EEG data. The dashed lines divide the segment to three parts: state 1, transition and state2. b) The concatenated parameter evolutions Φ_1^c, Φ_2^c and the corresponding smoothed evolutions. c) An adapted realization.

TABLE 2

The segment lengths T_i drawn from the distributions $g_1(T)$ and $g_2(T)$ and the predictor coefficients $\phi_j,\ j = 1, \ldots, 6$ and prediction error variances γ drawn from the normal distribution with means $\hat{\mu}_1$, $\hat{\mu}_2$ and covariances $\hat{C}_1$ and $\hat{C}_2$. The prediction error marked with an asterix has been constrained to lie in the observed region.

State	$\sum T_i$	T_i	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	ϕ_6	γ
1	0	645.5	0.3402	0.0455	-0.02277	-0.08406	0.0358	-0.2347	0.437
2	645	410.2	0.7642	-0.3862	-0.2483	0.06458	-0.01698	-0.3176	2.224
1	1056	804.8	0.25	0.161	-0.1543	0.06727	0.05747	-0.295	0.3∗
2	1861	532.5	0.9159	-0.3473	-0.53	0.254	0.106	-0.4673	1.159
1	2393	1368	0.4241	0.05525	0.03071	0.01801	0.1109	-0.3348	0.3509
2	3761	555.4	0.774	-0.4261	-0.3595	0.02872	0.1362	-0.5736	1.522

to avoid the abrupt changes in process characterictics discussed in Section 3.1. See [6] for an example of the simulation of chaotic EEG. Another method to achieve this is to generate a noise process with constant spectrum but with appropriate phase coherence. This can then be fed to the time-varying filter.

References

[1] A. Benveniste, M. Metivier, and P. Priouret. *Adaptive Algorithms and Stochastic Approximations*. Springer-Verlag, 1990.

[2] K. Blinowska and M. Malinowski. Non-linear and linear forecasting of the EEG time series. *Biol Cybern*, 66:159–165, 1991.

[3] G. Bodenstein and H. Praetorius. Pattern recognition of EEG by adaptive segmentation. In W. Perkins, editor, *Biomedical Computing*. Pitman Medical, 1977.

[4] P. Fenwick, P. Mitchie, J. Dollimore, and G. Fenton. Mathematical simulation of the electroencephalogram using an autoregressive series. *Bio-med. Comput.*, 2:281–307, 1971.

[5] P.B.C. Fenwick, P. Mitchie, J. Dollimore, and G.W. Fenton. Application of the autoregressive model to E.E.G. analysis. *Agressologie*, 10:553–564, 1969.

[6] W.J. Freeman. Simulation of chaotic EEG patterns with a dynamic model of the olfactory system. *Biol Cybern*, 56:139–150, 1987.

[7] W. Gersch. Spectral analysis of EEG's by autoregressive spectral decomposition of time series. *Math Biosci*, 7:205–222, 1970.

[8] W. Gersch. Non-stationary multichannel time series analysis. In *Methods of Analysis of Brain Electrical and Magnetic Signals*, volume 1 of *Handbook of Electroencephalography and Clinical Neurophysiology*, pages 261–296. Elsevier, 1987.

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

[10] A. Isaksson. On time variable properties of EEG signals examined by means of a Kalman filter. Technical Report 95, Royal Institute of Technology, 1975.

[11] N. Ishii, A. Iwata, and N. Suzumura. Segmentation of non-stationary time series. *Int J Systems Sci*, 10:883–894, 1979.

[12] J. Kaipio. Personal collection of papers on EEG modeling, 1994.

[13] B. Kemp and H.A.C. Kamphuisen. Simulation of human hypnograms using a Markov chain model. *Sleep*, 9:405–414, 1986.

[14] L. Ljung. *System Identification Toolbox for MATLAB*.

[15] L. Ljung. *System Identification: Theory for the User*. Prentice-Hall, 1987.

[16] O.M. Markand. Alpha rhythms. *J Clin Neurophysiol*, 7:163–189, 1990.

[17] L. Marple. A new autoregressive spectrum analysis algorithm. *IEEE Trans Acoust, Speech and Signal Processing*, 28:441–454, 1980.

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

[19] J. Pijn, J. van Veerven, A. Noest, and F. Lopez da Silva. Chaos or noise in EEG signals; dependence on state and brain site. *Electroencephalogr Clin Neurophysiol*, 79:371–381, 1991.

[20] 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*, pages 493–509. Raven Press, 1976.

[21] A. Wennberg and L. Zetterberg. Application of a computer-based model for EEG analysis. *Electroencephalogr Clin Neurophysiol*, 31:457–468, 1971.

[22] J. Wright, R. Kydd, and A Sergejev. Autoregression models of EEG. *Biol Cybern*, 62:201–210, 1990.

[23] L. Zetterberg. Estimation of parameters for a linear difference equation with application to EEG analysis. *Math Biosci*, 5:227–275, 1969.

[24] L. Zetterberg. Experience with analysis and simulation of EEG signals with parametric description of spectra. In P. Kellaway and I. Petersen, editors, *Automation of Clinical Electroencephalography*, pages 161–201. Raven Press, 1973.