Learning dynamics from nonstationary time series: Analysis of electroencephalograms

Dmitrii Gribkov* and Valentina Gribkova

Department of Physics, Lomonosov Moscow State University, 119899, Moscow, Russia

We propose an empirical modeling technique for a nonstationary time series analysis. Proposed methods include a high-dimensional ($N>3$) dynamical model construction in the form of delay differential equations, a nonparametric method of respective time delay calculation, the detection of quasistationary regions of the process by reccurence analysis in the space of model coefficients, and final fitting of the model to quasistationary segments of observed time series. We also demonstrate the effectiveness of our approach for nonstationary signal classification in the space of model coefficients. Applying the empirical modeling technique to electroencephalogram (EEG) records analysis, we find evidence of high-dimensional nonlinear dynamics in quasistationary EEG segments. Reccurence analysis of model parameters reveals long-term correlations in nonstationary EEG records. Using the dynamical model as a nonlinear filter, we find that different emotional states of subjects can be clearly distinguished in the space of model coefficients.

PACS number(s): 05.45.−a, 07.05.Tp, 07.05.Kf, 87.19.La

I. INTRODUCTION

Most observed real world dynamical processes are nonstationary. The nonstationarity coming from a changing environment essentially restricts the possibility of dynamics extraction from the corresponding time series. Dynamical invariants and stationary statistical characteristics of the process no longer have any sense in the absence of stationary measures.

However, the reality is not so pessimistic. Open systems demonstrate a quasistationary behavior if the time scales of their intrinsic motion are incomparable with characteristic times of external action, and if system parameters change adiabatically. Also, quasistationary segments of time series occur if the generating system is sensitive to resonant action only, or has a definite threshold of response to external forces. The existence of quasistationarity provides us with the possibility to reveal ''hidden'' dynamics from the observed process. However, detection of quasistationary regions in chaotic time series with high embedding dimensions is a complicated problem, and its solution depends on the type of nonstationarity, the length of the observed time series, the noise level, the dimensionality, and other factors. Generally, we cannot talk about the nonstationarity or stationarity of the process in the absence of an appropriate dynamical model of phenomena. Methods [1–6] which do not suggest a prior knowledge of the dynamical model of the process detect mostly obvious changes in observables, but fall into problems of detection of nonstationarities distorting correlations of the highest orders. Methods applying local dynamical models [7,6,8], or low-dimensional global models [9], have the same problems, since they can not reveal the long-term correlations and global dynamics of high-dimensional system. On the other hand, the construction of global empirical models governing the process is not a trivial task for high-dimensional chaotic dynamics manifesting itself in a scalar time series. Nevertheless, several attempts have been made recently towards solving this problem [10,11]. Occasionally, we can derive an empirical model which represents prominent features of the real dynamics under consideration. However, the correct solution to the problem depends not only on the length of the observed time series and level of its nonstationarity, but also on the ability of fitting model to take into account the highest dynamical degrees of freedom and to govern the long-term correlations of the process. Generally, we formulate the task of dynamical model construction from nonstationary scalar time series as consisting of four basic steps: (1) Revealing the basis of variables containing maximal information about the process. (2) Finding an appropriate model representation in terms of global maps or sets of differential equations. (3) Application of the model as a nonlinear filter for quasistationary segment detection. (4) Final fitting of the constructed model to quasistationary regions of the process.

In this paper we make an attempt to develop an approach which could provide us with an effective instrument of nonstationary time series analysis and processing. Since the subject of our analysis is human electroencephalogram (EEG) records, we address our work to those readers who are interested in EEG signal processing.

The paper consists of four main parts. In Sec. II we introduce the consequent decomposition of the state vector of the process with a high embedding dimension. Section III includes a high-dimensional dynamical model construction in terms of delay differential equations (DDEs) and a nonparametric estimation of the corresponding time delay. In Sec. IV we apply recurrence analysis [12–14,3,6] in the space of model coefficients for the purpose of quasistationary region detection in EEG records. We then fit the high-dimensional model to quasistationary regions of an EEG time series. Finally, in Sec. V we demonstrate the technique of emotional state classification in the space of model parameters derived from the nonlinear processing of respective EEG records.

II. STATE VECTOR DECOMPOSITION

In this section we begin with a linear separation of low- and high-dimensional components of the state vector of the

*Electronic address: dima_g@pacbell.net

observed process. The basic idea of the decomposition is to determine two sets of variables which independently describe different scales of motion. Let $x_k = x(t_k)$, $k = 1,...,N_t$ be a time series of the observed process $x(t)$. We define the state vector of the process as $\vec{R}_k = \vec{R}(x_{1k}, x_{2k}, ..., x_{Nk})$ in an N-dimensional embedding space, where $x_{ik} = x[t_k + (i - 1)\tau]$, and τ is the respective time step. The corresponding covariance matrix $\hat{C}$ of the process is given as

$$\hat{C} = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1N} \\ \vdots & & \vdots \\ \sigma_{N1} & \cdots & \sigma_{NN} \end{bmatrix}, \tag{1}$$

where $\sigma_{ll} = \Sigma_{k=1}^{N_t} x_{lk}^2$ and $\sigma_{jl} = \Sigma_{k=1}^{N_t} x_{jk} x_{lk}$. We then find eigenvectors $\vec{S}_i = \vec{S}(\mu_{i1}, ..., \mu_{iN})$, $i = 1, ..., n$ of matrix (1), solving

$$(\hat{C} - \lambda_i \hat{I})\vec{S}_i = 0, \tag{2}$$

where μ_{ij} is the jth coordinate of the ith eigenvector, λ_i is the ith eigenvalue, and $\hat{I}$ is the unit matrix. Equation (2) is a ''routine'' eigenvalue problem. Using a linear transform $\tilde{x}_i = (\vec{R}, \vec{S}_i)$ [15], we obtain new variables $\tilde{x}_i$, which are orthogonal in the sense of

$$\int_{-T}^{T} \tilde{x}_i \tilde{x}_j dt = 0 \tag{3}$$

for $T \to \infty$ and $i \neq j$. Actually, we can write scalar product in Eq. (3) as

$$\int_{-T}^{T} \tilde{x}_i \tilde{x}_j dt = \sum_{m=1}^{N} \sigma_{mm} \mu_{im} \mu_{jm} + \sum_{l \neq m, \, l,m=1}^{N} \sigma_{lm} \mu_{il} \mu_{jm}, \tag{4}$$

where $\sigma_{mm} = \int_{-T}^{T} x_m^2 dt$ and $\sigma_{lm} = \int_{-T}^{T} x_l x_m dt$. Using the vector form of Eq. (4), we obtain

$$\int_{-T}^{T} \tilde{x}_i \tilde{x}_j dt = \vec{S}_j^T \hat{C} \vec{S}_i. \tag{5}$$

Finally, from Eqs. (5) and (2) we have

$$\vec{S}_j^T \hat{C} \vec{S}_i = \lambda_i (\vec{S}_j^T, \vec{S}_i) = 0$$

due to the orthogonality of the two different eigenvectors. In order to reveal the structure of the new variables $\tilde{x}_i$, we choose, for simplicity, a number of eigenvectors $n = 3$ and an embedding dimension $N = 3$. Taking into account $\sigma_{11} = \sigma_{22} = \sigma_{33} = \sigma_0 = $ const, for $T \to \infty$, from Eq. (4) we obtain

$$\int_{-T}^{T} \tilde{x}_i \tilde{x}_j dt = (\mu_{i1}\mu_{j2} + \mu_{i2}\mu_{j1})\sigma_{12} + (\mu_{i1}\mu_{j3} + \mu_{i3}\mu_{j1})\sigma_{13}$$

$$+ (\mu_{i2}\mu_{j3} + \mu_{i3}\mu_{j2})\sigma_{23}. \tag{6}$$

Since $\sigma_{12} = \sigma_{23}$ for $T \to \infty$, using Eqs. (6) and (3), we have

$$(\mu_{i1}\mu_{j2} + \mu_{i2}\mu_{j1} + \mu_{i2}\mu_{j3} + \mu_{i3}\mu_{j2})\sigma_{12} + (\mu_{i1}\mu_{j3}$$

$$+ \mu_{i3}\mu_{j1})\sigma_{13} = 0. \tag{7}$$

Adding conditions of orthogonality and normalization of eigenvectors, we find, from Eqs. (7) and (3):

$$\mu_{i1}\mu_{j2} + \mu_{i2}\mu_{j1} + \mu_{i2}\mu_{j3} + \mu_{i3}\mu_{j2} = 0,$$

$$\mu_{i1}\mu_{j3} + \mu_{i3}\mu_{j1} = 0,$$

$$\mu_{i1}\mu_{j1} + \mu_{i2}\mu_{j2} + \mu_{i3}\mu_{j3} = 0, \tag{8}$$

$$\mu_{i1}^2 + \mu_{i2}^2 + \mu_{i3}^2 = 1,$$

$$\mu_{j1}^2 + \mu_{j2}^2 + \mu_{j3}^2 = 1.$$

Solving Eqs. (8), we obtain three eigenvectors

$$\vec{S}_1 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right),$$

$$\vec{S}_2 = \left(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right),$$

$$\vec{S}_3 = \left(\frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}} \right).$$

New variables $\tilde{x}_{ik} = \tilde{x}_i(t_k)$ follow from the scalar product $\tilde{x}_{ik} = (\vec{R}_k, \vec{S}_i)$:

$$\tilde{x}_{1k} = \frac{1}{\sqrt{3}} \sum_{j=1}^{3} x_{jk},$$

$$\tilde{x}_{2k} = \frac{1}{\sqrt{2}} [x_{3k} - x_{1k}], \tag{9}$$

$$\tilde{x}_{3k} = \frac{1}{\sqrt{6}} [x_{1k} - 2x_{2k} + x_{3k}],$$

where $\tilde{x}_{2k}$ and $\tilde{x}_{3k}$ are proportional to first and second order derivatives of the process, respectively. Variables (9) determine the new low-dimensional ($n = 3$) state vector in the space of the eigenvectors;

$$\vec{r}_k = \sum_{i=1}^{n} \vec{S}_i \tilde{x}_{ik}. \tag{10}$$

We can rewrite Eq. (10) in terms of the original N-dimensional embedding space

$$\vec{r}_k = \sum_{j=1}^{N} \vec{i}_j x_{jk}^r, \tag{11}$$

where $\vec{i}_j$ is the jth basis vector, and x_{jk}^r is the jth coordinate of the low-dimensional state vector $\vec{r}_k$ embedded into the original N-dimensional space. Comparing Eqs. (10) and (11), we find

$$\vec{r}_k = \sum_{j=1}^{N} \vec{i}_j \sum_{i=1}^{n} \mu_{ij} \tilde{x}_{ik} = \sum_{j=1}^{N} \vec{i}_j \sum_{i=1}^{n} \mu_{ij} \sum_{m=1}^{N} \mu_{im} x_{mk}. \tag{12}$$
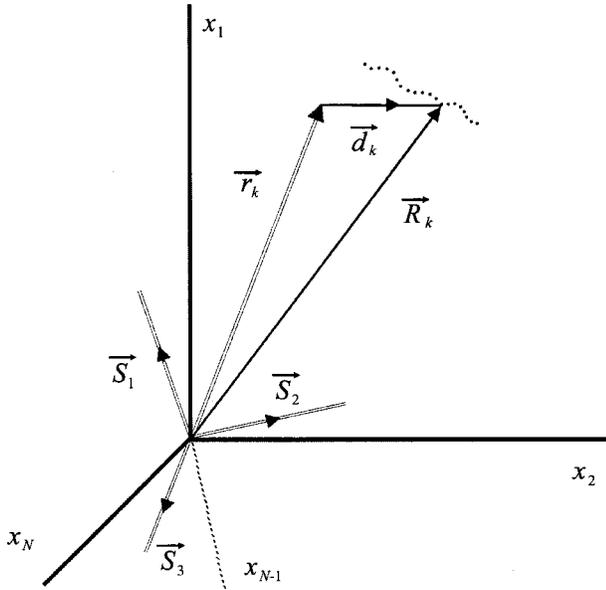
FIG. 1. State vector $\vec{R}_k$ decomposition: $\vec{r}_k$ is the low-dimensional component defined by eigenvectors $\vec{S}_i$, and $\vec{d}_k$ is the vector of difference accounting for higher degrees of freedom.

Formula (12) defines the low-dimensional vector $\vec{r}_k$ in the N-dimensional embedding space. $\vec{r}_k$ does not contain all the information about the process. If the minimal dimension of the embedding space $N_{min} > n$, we define the vector $\vec{d}_k = \vec{R}_k - \vec{r}_k$, which is the difference between the original N-dimensional state vector $\vec{R}_k$ and the low-dimensional $\vec{r}_k$ (Fig. 1). Using Eq. (12), we obtain

$$\vec{d}_k = \sum_{j=1}^{N} \vec{i}_j x_{jk}^d, \tag{13}$$

$$x_{jk}^d = x_{jk} - \sum_{i=1}^{n} \mu_{ij} \sum_{m=1}^{N} \mu_{im} x_{mk}. \tag{14}$$

We suggest that vector $\vec{d}_k$ includes information about the highest degrees of freedom which are not accounted for by $\vec{r}_k$.

III. HIGH-DIMENSIONAL DYNAMICAL MODEL CONSTRUCTION FOR STATIONARY TIME SERIES

In Sec. II we did not regard the nonstationarity of the observed time series. This does not play a significant role in the formal decomposition of state vector $\vec{R}_k$. We only made the assumption that vector $\vec{d}_k = \vec{R}_k - \vec{r}_k$ includes information about the highest degrees of freedom which are not accounted for by state vector $\vec{r}_k$, responsible for the low-dimensional dynamics derived from the observed time series. However, in this section we use stationary measures, and must require the process to be stationary.

Since the second and third variables in Eq. (9) approximate the first and second order derivatives of the process, respectively, it is reasonable to present the dynamics of the low-dimensional vector $\vec{r}(t)$ as

$$\frac{dx}{dt} = y, \quad \frac{dy}{dt} = z, \quad \frac{dz}{dt} = f(x,y,z,\vec{a}), \tag{15}$$

where variable x represents an observable, and f and $\vec{a}$ are the nonlinearity and the vector of the coefficients, correspondingly. We choose f in the form of a polynomial of Lth order. Unfortunately, high levels of noise in real processes do not allow us to apply model representation (15) to $\vec{d}_k$. We cannot use Eq. (14) as a local variable, because vector $\vec{d}_k$ contains high order derivatives that produce large errors. However, we can take into account the highest degrees of freedom "hidden" in Eqs. (13) and (14), implicitly.

We consider an expansion of system (15) in a class of DDE,

$$\frac{dx}{dt} = y, \quad \frac{dy}{dt} = z, \quad \frac{dz}{dt} = f(x,y,z,x_\tau,\vec{a}), \tag{16}$$

where $x_\tau = x(t - \tau^*)$, and τ^* is a characteristic time delay. As always, we have a problem of how closely the constructed model approximates the dynamics of our observable. We choose function $f = \sum_{m=1}^{M} a_m \varphi_m$ in the form of a polynomial of third order, with $M = 35$ nonlinear terms φ_m. In general, Eqs. (16) have infinite dimensions, and we can expect that system (16) will fit the high-dimensional dynamics of the observed time series. Another problem is how to estimate the value of τ^*. Since methods proposed in Ref. [10] use special forms of the nonlinearity f, and do not work for polynomials of general form, we apply another nonparametric approach to time delay estimation. We look for a delay variable satisfying the following requirement: x_τ is statistically independent of x, and contains maximal information about vector $\vec{d}_k$. This requirement leads to calculations of extrema of two pieces of mutual information:

$$I_\alpha(\tau^*) = \sum_{i,j} p(x_i, x_{\tau j}) \log \left[\frac{p(x_i, x_{\tau j})}{p(x_i) p(x_{\tau j})} \right] \tag{17}$$

and

$$I_\beta(\tau^*) = \sum_{i,j} p(x_i^d, x_{\tau j}) \log \left[\frac{p(x_i^d, x_{\tau j})}{p(x_i^d) p(x_{\tau j})} \right] \tag{18}$$

where $p(x_i, x_{rj})$ and $p(x_i)$ are joint and individual probability densities, correspondingly.

Functions $I_\alpha(\tau^*)$ and $I_\beta(\tau^*)$ have several extrema on a limited interval. We look for the minimum of Eq. (17) and the maximum of Eq. (18), which are close enough on some interval $[\tau_1, \tau_2]$. Thus we obtain two estimations τ_α and τ_β of time delay τ^*, respectively. In general, when values τ_α and τ_β are significantly different, we can construct a multi-delay dynamical system or, at least, try each value of the delay, using Eq. (16). As shown in the next sections, even the single delay model is sufficient for performing an EEG records analysis.

In some cases formula (18) can be replaced by a more simple relation. We find the maximal average of the absolute value $|x_{jk}^d|$ [see Eq. (14)],

$$\langle |x_j^d| \rangle_T = \frac{1}{N_t} \sum_{k=1}^{N_t} |x_{jk}^d|,$$

where T is the time of observation. The maximal average projection of $\vec{d}_k$ gives an index j^* of the delay coordinate in N-dimensional embedding space. We then find the delay as $\tau^* = j^* \Delta t$, where Δt is a sampling time interval of the time series.

We cannot give a strict recommendation as to the application of a certain formula or certain time delay in all possible cases. Each time series requires individual consideration. According to our experience, a prior approach and a definite freedom of choice of τ^* are more effective than parametric methods of delay estimation.

Having delay variable x_τ, we fit system (16) to the observed time series x_k by the solution of the matrix equation

$$\Delta \vec{z} = 2\Delta t \cdot \hat{\Phi}(\vec{x}, \vec{y}, \vec{z}, \vec{x}_\tau) \cdot \vec{a}, \qquad (19)$$

where

$$\hat{\Phi} = \begin{bmatrix} 1 & x_2 & y_2 & z_2 & x_{\tau 2} & x_2^2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{N_t-1} & y_{N_t-1} & z_{N_t-1} & x_{\tau N_t-1} & x_{N_t-1}^2 & \cdots \end{bmatrix}$$

is a matrix of nonlinearities corresponding to f, and

$$\Delta \vec{z} = \begin{bmatrix} z_3 - z_1 \\ \vdots \\ z_{N_t} - z_{N_t-2} \end{bmatrix}, \quad \vec{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_M \end{bmatrix}.$$

Using Eq. (19), we can obtain vector $\vec{a}$ of the model coefficients from

$$\hat{\Phi}^T \cdot \Delta \vec{z} = 2\Delta t \cdot \hat{\Phi}^T \cdot \hat{\Phi} \cdot \vec{a},$$

where $\hat{\Phi}^T$ is a transposed matrix $\hat{\Phi}$. Having this set of fitted coefficients, we integrate Eqs. (16) numerically by the simplest Euler method adapted to DDE integration, with initial conditions x_1, y_1, z_1, and $x_\tau = x_{1-n^*}$ derived from observed time series.

IV. QUASISTATIONARY REGION DETECTION AND MODELING

Next we suppose that our observable is nonstationary. Our main goal is to determine regions where the time series demonstrates a quasistationary behavior. In other words, we want to find regions of long-term correlations and similar dynamics (if the latter is present). We define a sampling time window of length n_w moving along the process, and estimate the time delay in each window, using formulas (17) and (18). We then fit model (16) to the data bounded by each moving window. Equations (16) are assumed to be a sort of nonlinear filter matching the observed time series. The length n_w of the window is chosen to be proportional to the characteristic time of the process of correlation distortion. Each set of fitted coefficients is represented by an M-dimensional vector $\vec{a}_i (i=1, \ldots, N_w)$, where $M = 35$, i is the index of the current window, and $N_w \simeq N_t - n_w$ is the whole number of sampling
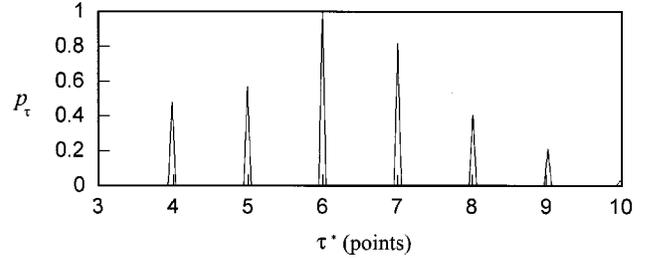
FIG. 2. Probability distribution p_τ vs time delay τ^* for a quasistationary segment of the EEG record registered from a subject in the rest state.

windows. In order to reveal the dynamical correlations, we use the recurrence analysis technique in the space of model coefficients. However, some nonlinear terms of the model may be erroneous and add only numerical noise. Such terms can alter the dynamics of the original system. If the respective nonlinearity does not fit the time series, its coefficient is small and the variance of the coefficient (computed from the least-squares fit) is large. We can reduce the influence of erroneous nonlinearities, replacing each vector $\vec{a}_i = (a_{1i}, \ldots, a_{Mi})$ by vector $\vec{c}_i = (a_{1i}/|\delta a_{1i}|, \ldots, a_{Mi}/|\delta a_{Mi}|)$, where $|\delta a_{mi}|$ is the absolute value of the variance of the mth coefficient in the ith sampling window. We define Euclidean distance between two arbitrary chosen vectors $\vec{c}_i$ and $\vec{c}_j$, as $\Delta_{ij} = \|\vec{c}_i - \vec{c}_j\|$. If Δ_{ij} is less than a prior constant ε, we mark the point with coordinates (i,j) on the corresponding recurrence plane. Processing all $N_w \times N_w$ points, we obtain the recurrence plot.
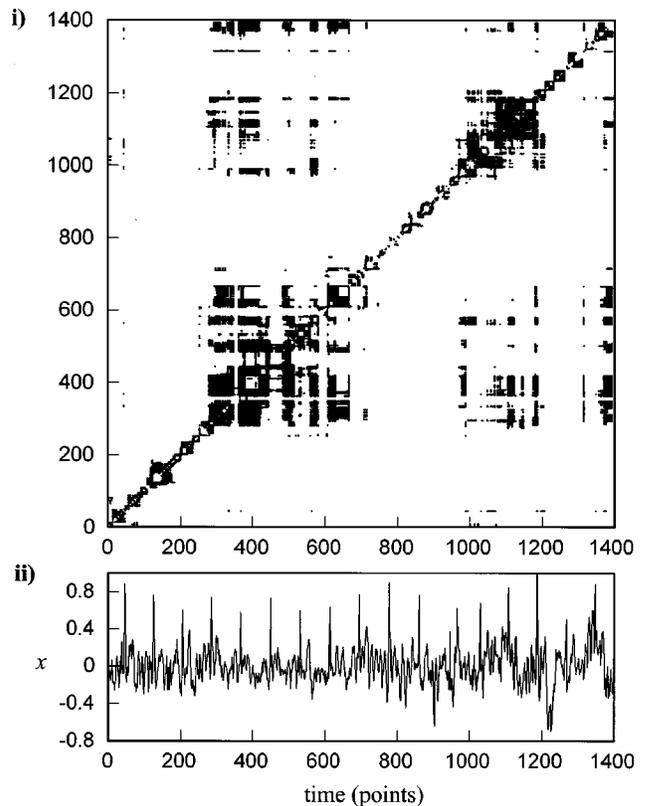
FIG. 3. (i) Reccurence plot of model (16) vectors $\vec{c}_i$, tracing the normalized EEG observable presented in (ii). Sets of black points mark regions of maximal correlations of coefficients.

TABLE I. Model (16) fitted to a quasistationary EEG segment.

m	φ_m	a_m	m	φ_m	a_m	m	φ_m	a_m
1	1.00	0.0	13	z^2	0.0	25	xx_τ^2	$-0.005\,62$
2	x	0.000 38	14	zx_τ	0.168 20	26	y^3	$-7.984\,41$
3	y	$-0.015\,88$	15	x_τ^2	0.0	27	y^2z	19.261 95
4	z	0.0	16	x^3	$-0.007\,05$	28	y^2x_τ	$-0.684\,71$
5	x_τ	$-0.001\,03$	17	x^2y	$-0.066\,27$	29	yz^2	-121.763
6	x^2	$-0.002\,875$	18	x^2z	$-0.667\,33$	30	yzx_τ	0.0
7	xy	0.054 87	19	x^2x_τ	0.0	31	yx_τ^2	0.0
8	xz	$-0.220\,58$	20	xy^2	0.0	32	z^3	-300.343
9	xx_τ	$-0.002\,41$	21	xyz	$-3.344\,29$	33	z^2x_τ	0.0
10	y^2	0.200 41	22	xyx_τ	$-0.055\,40$	34	zx_τ^2	0.0
11	yz	2.913 92	23	xz^2	-31.1660	35	x_τ^3	0.001 74
12	yx_τ	0.0	24	xzx_τ	0.273 42			

For demonstration of described technique we choose ''normal'' human EEG record with sampling frequency 200 Hz. The whole length of the time series is $N_t = 1600$ points. We choose window of length $n_w = 200$ points, moving one point ahead along the EEG observable. By fitting the model, we obtain $N_w = 1400$ vectors $\vec{c}_i$. Figure 2 demonstrates the distribution of delays τ_i^* calculated for each time window. Figure 3(i) presents recurrence plot of all pairs of normalized vectors $\vec{c}_i / \|\vec{c}_{max}\|$ with distances $\Delta_{ij} < 2^{-4}$. Sets of points in Fig. 3(i) determine regions of maximal correlations of model coefficients. The corresponding time series of the EEG record is plotted in Fig. 3(ii). We can expect to find a stationary dynamics which is governed by model (16) inside regions of strong correlations of model coefficients. Before fitting the model to observed data, we apply a five-point local spline approximation of the quasistationary segment (points 280–800) detected in Fig. 3(i). We then fit model (16) to the segment, from point 301 to point 500 (totally 10^3 points of interpolated data). The fitting parameters of the model are presented in Table I. Time delay τ^* is equal to 25 ($\tau^* = n^*\Delta t$, $n^* = 125$). While model construction all erroneous nonlinearities with small coefficients and large variances of coefficients are zeroed.

Having fitted coefficients, we integrate equations (16) forward in time 10^4 time points ahead with step $\Delta t = 0.2$, and initial conditions derived from real (interpolated) EEG observable. Because we apply a fractional time step, the integration scheme provides additional interpolation of time series, in accordance with step division. Model (16) generates the time series which covers an unfitted interval of real data. The generated process is then compared with a real EEG segment, from point 351 to point 750 (totally 2×10^3 points of interpolated data). Figures 4(i) and 4(ii) demonstrate phase trajectories and time series of real (interpolated) and model processes, respectively. We find that phase portraits of reconstructed and model trajectories are topologically similar. However, estimations of embedding dimensions of real and model time series are different and equal to 8 and 6, correspondingly. Nevertheless, in contrast to local modeling techniques, we obtain a global model generating the process similar to the quasistationary segment of observable. The advantage of such technique is that we can match the long-term correlations of the process.
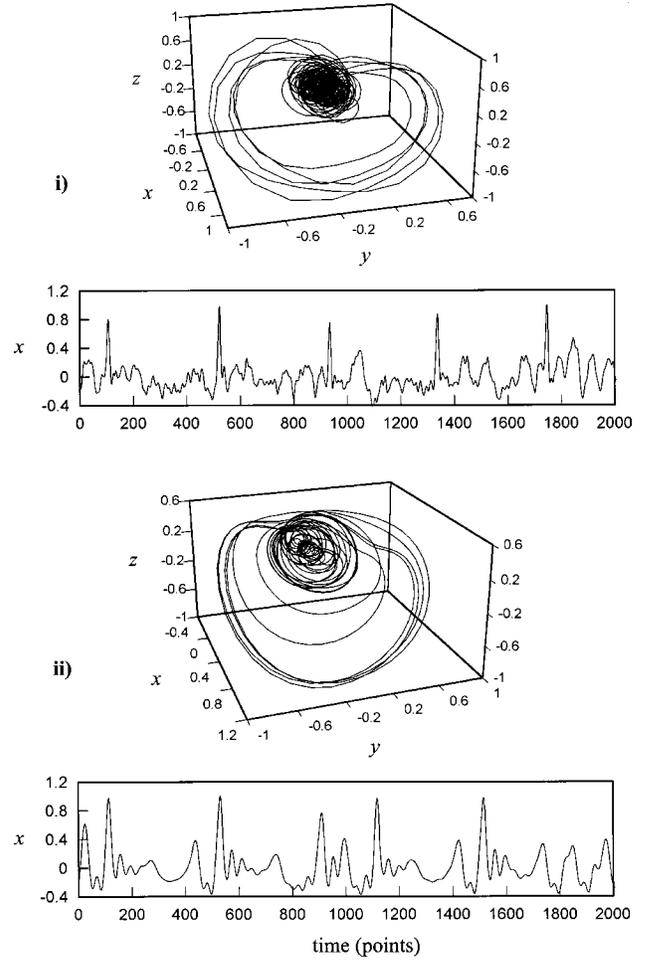
FIG. 4. (i) Phase portrait and time series of a real quasistationary EEG segment. (ii) Trajectory and time series generated by model (16), with parameters listed in Table I.

The next example demonstrates the empirical models constructed for two quasistationary segments revealed in a single EEG record registered from the subject, whose eyes were opened and closed during some period. The sampling frequency of the EEG record is 128 Hz. Quasistationary segments for the ''open eyes'' state (I) and the ''closed eyes'' state (II) are detected by the recurrence technique described

TABLE II. Model (16) fitted to a quasistationary EEG segment (state I).

m	φ_m	a_m	m	φ_m	a_m	m	φ_m	a_m
1	1.00	0.0	13	z^2	$-3.274\,86$	25	xx_τ^2	0.0
2	x	0.0	14	zx_τ	0.0	26	y^3	0.0
3	y	$-0.063\,34$	15	x_τ^2	0.0	27	y^2z	-12.454
4	z	$-0.030\,94$	16	x^3	$-0.008\,30$	28	y^2x_τ	0.0
5	x_τ	$-0.035\,34$	17	x^2y	$-0.066\,63$	29	yz^2	0.0
6	x^2	$-0.009\,35$	18	x^2z	$-0.258\,36$	30	yzx_τ	-4.1356
7	xy	$-0.032\,64$	19	x^2x_τ	0.005 43	31	yx_τ^2	$-0.218\,49$
8	xz	$-0.336\,62$	20	xy^2	$-0.238\,62$	32	z^3	-99.525
9	xx_τ	0.002 69	21	xyz	0.0	33	z^2x_τ	0.0
10	y^2	0.117 52	22	xyx_τ	$-0.137\,83$	34	zx_τ^2	0.0
11	yz	$-0.820\,27$	23	xz^2	-6.2304	35	x_τ^3	$-0.007\,83$
12	yx_τ	$-0.062\,83$	24	xzx_τ	0.092 00			

TABLE III. Model (16) fitted to a quasistationary EEG segment (state II).

m	φ_m	a_m	m	φ_m	a_m	m	φ_m	a_m
1	1.00	0.0	13	z^2	0.0	25	xx_τ^2	0.0
2	x	0.0	14	zx_τ	0.113 69	26	y^3	$-3.656\,61$
3	y	$-0.021\,76$	15	x_τ^2	$-0.001\,67$	27	y^2z	-7.8836
4	z	0.058 57	16	x^3	0.000 66	28	y^2x_τ	$-0.849\,87$
5	x_τ	$-0.001\,74$	17	x^2y	$-0.017\,66$	29	yz^2	161.65
6	x^2	$-0.000\,21$	18	x^2z	0.0	30	yzx_τ	0.0
7	xy	0.011 925	19	x^2x_τ	$-0.000\,77$	31	yx_τ^2	$-0.099\,87$
8	xz	0.0	20	xy^2	0.0	32	z^3	-328.63
9	xx_τ	0.001 60	21	xyz	3.084 17	33	z^2x_τ	16.1856
10	y^2	0.0	22	xyx_τ	$-0.008\,01$	34	zx_τ^2	0.0
11	yz	0.783 63	23	xz^2	-9.7743	35	x_τ^3	$-0.003\,34$
12	yx_τ	$-0.017\,68$	24	xzx_τ	0.359 23			

above. Here we also apply a five-point local spline interpolation of real data. Tables II and III present models fitted to quasistationary regions of states I and II, respectively. Model I is found from 8×10^2 points, and model II obtained from 10^3 points of interpolated data. Delays calculated for states I and II are equal to 70 and 18, correspondingly. We integrate
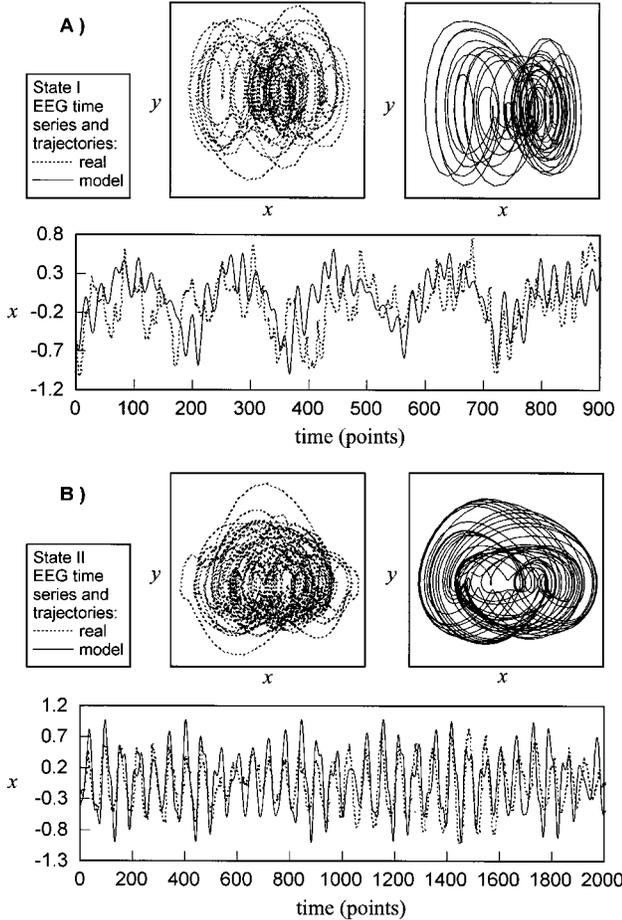
FIG. 5. (A) Trajectories and time series of real and model quasistationary EEG segments for state I (open eyes). (B) Trajectories and time series of real and model quasistationary EEG segments for state II (closed eyes).
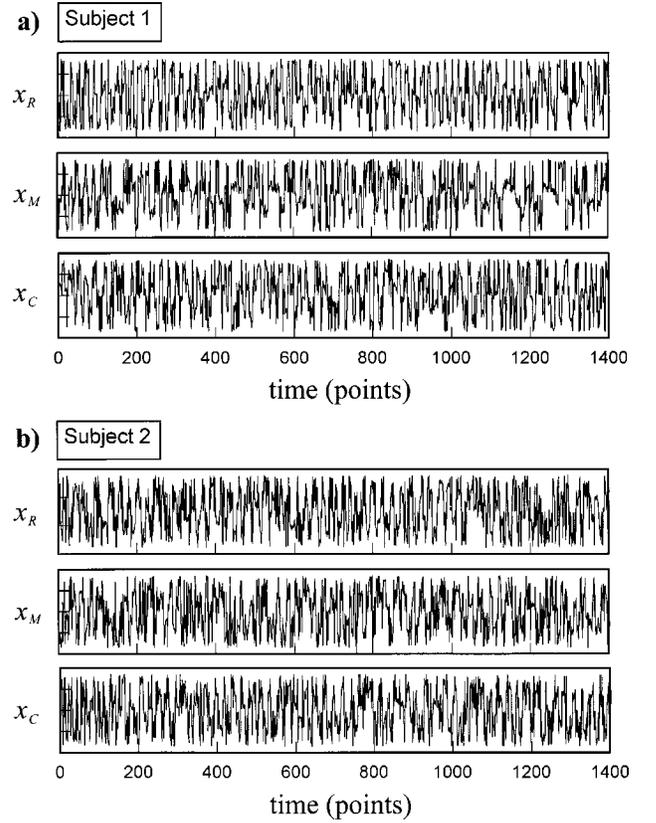
FIG. 6. Normalized EEG records x_R, x_M, and x_C registered from subject 1 (a) and subject 2 (b) in R, M, and C emotional states, respectively.

model I (Table II) 2700 points ahead with step $\Delta t = 1/3$. Figure 5(A) demonstrates real (interpolated) and model time series and trajectories for state I. Model II (Table III) is integrated 5×10^4 points ahead with step $\Delta t = 4 \times 10^{-2}$. The respective time series and trajectories for real and model processes are plotted in Fig. 5(B). Generally, the contrast between states I and II is obvious and does not require a complicated modeling technique, but in Sec. V we consider a classification of emotional states where dynamical discrepancies can be fixed only by changes of model coefficients.

V. CLASSIFICATION IN THE SPACE OF COEFFICIENTS

Here we demonstrate the technique of emotional state classification in the space of model coefficients. We consider EEG records of three emotional states: a ''rest'' (R) state with closed eyes, a ''listening music'' (M) state, and a ''calculation'' (C) state. A subject in the M state was requested to listen to music with closed eyes, and a subject in the C state was asked to calculate in his mind with eyes closed. EEG records were registered from different subjects with a sampling frequency of 200 Hz. Each time series includes 1400 points (7 sec) of real EEG records [Figs. 6(a) and 6(b)]. Using a five-point spline interpolation of observed data, we obtain about 7×10^3 points. Applying sampling and fitting procedures described in Sec. IV, we obtain a set of delays and model vectors $\vec{a}_i$. The parameters are calculated in windows of length $n_w = 10^3$ points. In order to avoid local correlations induced by the spline approximation, we choose a step of window movement equal to ten points. Finally, we obtain $N_w = 6 \times 10^2$ vectors $\vec{a}_i$ of the coefficients. For a clear
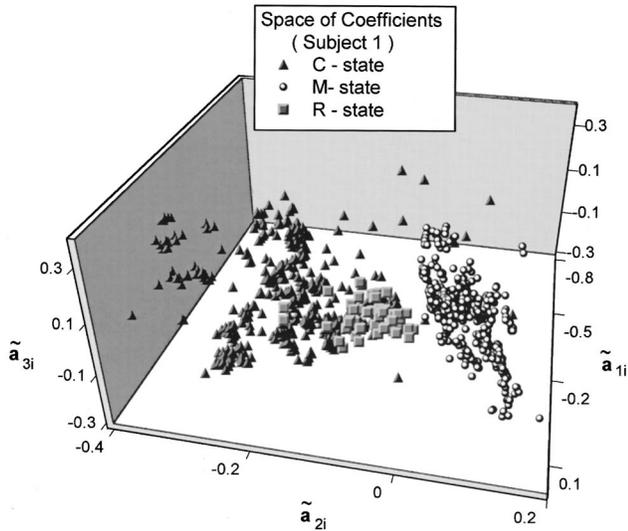
FIG. 7. Sets of points characterizing three emotional states of subject 1 in the space of model parameters $\tilde{a}_{li}$ ($l = 1, 2,$ and 3; $i = 1, \ldots, 600$).

presentation of the classification procedure, we consider the eigenvalue problem

$$(\hat{P} - \theta_l \hat{I})\vec{e}_l = 0,$$

$$\hat{P} = \hat{A}^T \hat{A}, \quad \hat{A} = \begin{bmatrix} a_{11} & \cdots & a_{1N_w} \\ \vdots & & \vdots \\ a_{M1} & \cdots & a_{MN_w} \end{bmatrix}, \quad (20)$$

where a_{ji} is the jth fitted coefficient in the ith sampling window, and θ_l and $\vec{e}_l$ are the lth eigenvalue and an eigenvector of matrix $\hat{P}$, respectively. For graphic simplicity we take $l = 3$. Solving Eq. (20), we obtain new parameters $\tilde{a}_{li} = (\vec{a}_i, \vec{e}_l)$ corresponding to the largest eigenvalues of matrix $\hat{P}$. Coordinates $\tilde{a}_{li}$ determine the evolution of a new vector $\vec{\alpha}_i = (\tilde{a}_{1i}, \tilde{a}_{2i}, \tilde{a}_{3i})$ in three-dimensional space. Figure 7 shows distinct separation of three sets of coefficients $\tilde{a}_{li}$ for three emotional states of subject 1. Figure 8 demonstrates a classification of M and C states of subjects 1 and 2 in a space of parameters $\tilde{a}_{li}$. We can see the intersections of sets of points characterizing two identical emotional states, and a clear distinction of the sets corresponding to different states. We check this interesting fact on two other subjects, and obtain similar results. Of course, our statistical ensemble (four subjects) is far from being sufficient to make any conclusion about invariant properties of distributions of vectors $\vec{\alpha}_i$ for various emotional states of different subjects; nevertheless, this interesting fact attracts our attention, and requires additional treatment. Here we only note the fact that the classification scheme works well for nonstationary signals with high embedding dimensions.
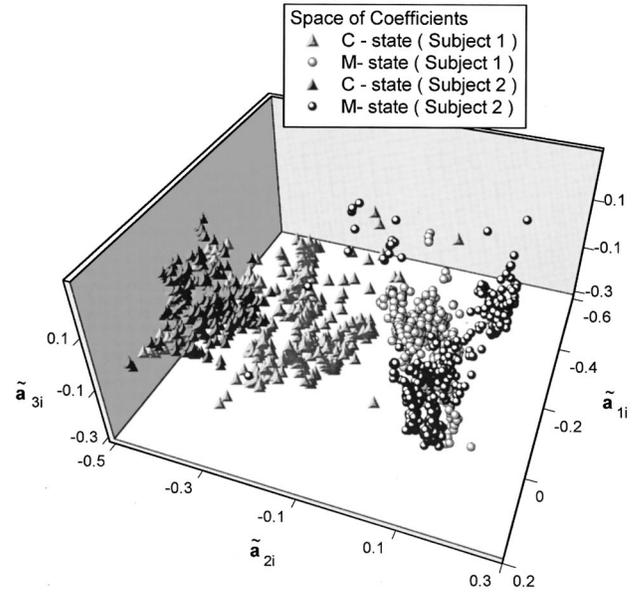
FIG. 8. Sets of model parameters $\tilde{a}_{li}$, ($l = 1, 2,$ and 3; $i = 1, \ldots, 600$) characterizing (M) and (C) emotional states of two subjects.

VI. CONCLUSION

A nonstationary dynamical time series analysis suggests the revelation of long- term nonlinear correlations and quasistationary regions of the process. These can be found via the construction of an empirical dynamical model of the observed process. Since most real observables have a high embedding dimension, the empirical model must be also high dimensional. Because there is a problem in finding an optimal high-dimensional basis of local variables from scalar observables, we apply an extended representation of the model in a class of DDE's. Thus our main results are as follows: (1) The construction of model basis consisting of orthogonal local variables which determine the evolution the of low- dimensional component of a state vector, and a statistically independent delay variable(s) which implicitly represents the high- dimensional components. (2) High- dimensional empirical model construction in the form of delay equations. (3) Application of recurrence analysis in the space of model parameters for quasistationary region detection. (4) A classification procedure defined in the space of model coefficients for nonstationary time series classification. The proposed technique demonstrates its efficiency for EEG record processing and classification.

ACKNOWLEDGMENTS

[1] M. B. Priestley, *Non-linear and Non-stationary Time Series Analysis* (Academic Press, London, 1988).

[2] H. Isliker, and J. Kurths, Int. J. Bifurcation Chaos Appl. Sci. Eng. **3**, 1573 (1993).

[3] C. L. Webber, and J. P. Zbilut, J. Appl. Physiol. **76**, 965 (1994).

[4] J. J. Zebrowski, W. Poplawska, and R. Baranowski, Acta Phys. Pol. B **26**, 1055 (1995).

[5] M. B. Kennel, Phys. Rev. E **56**, 316 (1997).

[6] T. Schreiber, Phys. Rep. **308**, 1 (1999); **2**, 3082 (1999).

[7] T. Schreiber, Phys. Rev. Lett. **78**, 843 (1997).

[8] O. Anosov *et al.*, in *Proceedings of the of the International Conference on Dynamical Systems and Chaos, Tokyo*, *May 1994* (World Scientific, Singapore, 1994), Vol. 2, p. 370.

[9] R. Brown, E. R. Rulkov, and N. F. Tracy, Phys. Rev. E **49**, 3784 (1994).

[10] S. P. Ellner, *et al.*, Physica D **110**, 182 (1997).

[11] H. U. Voss, A. Schwache, J. Kurths, and F. Mitschke, Phys. Lett. A **256**, 47 (1999).

[12] J. P. Eckmann, S. O. Kamphorst, and D. Ruelle, Europhys. Lett. **4**, 973 (1987).

[13] M. Kobbe and G. Mayer-Kress, in *Santa Fe Institute Studies in the Science of Complexity, Proceedings Vol. XII* (Addison-Wesley, Reading, MA, 1992).

[14] M. Gasdagli, Physica D **108**, 206 (1997).

[15] D. S. Broomhead and G. P. King, Physica D **20**, 217 (1986).