Identification of continuous–discrete state space models with delays

J.C. Jimenez*and T. Ozaki†

December 1997

Abstract

This report deals with the identification of continuous-discrete state space models with delay, that is, given a time series generated by this kind of model, to estimate of both the unknown system parmeters (including the delays) and the non-observable state variables. In particular, an identication method based on the Local Linearization filter and the maximum likelihood innovation approach is considered.

1 Introduction

This report deals with the identification of the Delay State Space (DSS) model defined by the continuous state equation

$$dx(t) = f(t, x(t), x(t-\tau_1), .., x(t-\tau_d), \theta)dt + \sigma(t, \theta)dW(t), \text{ for } t \geq t_0 \quad (1)$$

$$x(t) = \varphi(t), \text{ for -max}_{i=1..d}\{\tau_i\} \leq t \leq t_0 \quad (2)$$

and the discrete measurement equation

$$z_{t_k} = h(t_k, x(t_k), \theta) + e_{t_k}, \quad (3)$$

where $x(t) \in R^n$ is the state vector, $\sigma(t,\theta) \in R^{n\times m}$, W is an m-dimensional Winner process, $\tau = \{\tau_1, .., \tau_d\}$ is a set of delay parameters, $\theta = \{\theta_1, .., \theta_p\}$ is a set of other parameters, $z_{t_k} \in R^r$ is the observation vector, and $e_{t_k} \sim$

*Centro de Neurociencias de Cuba, Apartado 6880, La Habana, Cuba.

†Institute of Statistical Mathematics, 4-6-7 Minami Azabu, Minato-ku, Tokyo 106, Japan

$N(0,\Sigma(\theta))$ i.i.d. vector of random variables for every $t_k = t_0 + k\Delta$, with $\Delta > 0$ and $k = 0, .., N$. Here it is assumed that: 1) τ and θ are non-random sets, 2) φ is a deterministic function, and 3) the delayed stochastic differential equation (1)-(2) has a unique solution (see [17]).

The identification problem to consider here is the following: *given the delay state space model (1)-(3) and a set of discrete observations* $\{z_{t_0}, .., z_{t_N}\}$*, to estimate both the non-observable components of the state vector* x *at the instants of time* $t_0, .., t_N$ *and the unknown sets of parameters* θ *and* τ. In this report some identification methods for the DSS model (1)-(3), based on the Local Linearization (LL) filters and a maximum likelihood innovation approach, are considered. These methods involve different LL filtering algorithms which are derived from different approximations of the well known extended Kalman filter. A main part of such approximations is strongly connected with LL schemes for the numerical integration of delayed differential equation. For these reasons, the two first sections deal with the LL schemes for delayed differential equations and the LL filtering algorithms. After that, in section 4, the identification methods are presented. To evaluate the performance of any one of these identification methods, given a time series generated by the DSS model (1)-(3), a numerical scheme for the integration of the delayed stochastic differential equation (1)-(2) is required. Therefore, in the section 5 some LL schemes for the numerical integration of the delayed SDEs are presented and, finally, a preliminary simulation study is presented in section 6.

2 Local Linearization schemes for the numerical integration of delay differential equations

Let the Delayed Differential Equation (DDE)

$$\begin{aligned} dx(t) &= f(t, x(t), x(t-\tau_1), .., x(t-\tau_d))dt, \text{ for } t \geq t_0, & (4) \\ x(t) &= \varphi(t), \text{ for -max}_{i=1..d}\{\tau_i\} \leq t \leq t_0 & (5) \end{aligned}$$

where $x(t) \in R^n$. The above equation can be approximated by the following Ordinary Differential Equation (ODE)

$$dx(t) = f(t, x(t), \Pi_1(t), .., \Pi_d(t))dt, \text{ for } t \geq t_0, \tag{6}$$

where

$$\Pi_i(t) = \left\{ \begin{array}{l} \varphi(t), \ \mathit{for} \text{ -max}_{i=1..d}\{\tau_i\} \leq t \leq t_0 \\ \mathit{an\ approximation\ to}\ x(t-\tau_i), \ \mathit{for}\ t \geq t_0 \end{array} \right\} \tag{7}$$

for $i = 1, .., d$.

The ODE (6) is an SDE with zero diffusion coefficient. Therefore, a LL scheme for SDEs can be used for the numerical integration of that ODE and in this way obtained a numerical scheme for the DDE (4)-(5). Thus, setting the diffusion coefficient equal to zero in any one of the LL schemes for non-autonomous SDEs [2, 12, 24] the following LL scheme for DDE is obtained:

$$\begin{aligned} x_{t_{k+1}} &= x_{t_k} + r_0(F_x(t_k, x_{t_k}; \tau), \Delta) F(t_k, x_{t_k}; \tau) \\ &\quad + \{\Delta r_0(F_x(t_k, x_{t_k}; \tau), \Delta) - r_1(F_x(t_k, x_{t_k}; \tau), \Delta)\} F_t(t_k, x_{t_k}; \tau) \end{aligned} \tag{8}$$

where $t_k = t_0 + k\Delta$ for $k = 0, 1, ..,$ Δ is the step-size of the scheme, $\tau = \{\tau_1, .., \tau_d\}$, $F(t, x; \tau) = f(t, x, \Pi_1(t), .., \Pi_d(t))$,

$$F_x(t, x; \tau) = \frac{\partial f}{\partial x}(t, x, \Pi_1(t), .., \Pi_d(t)),$$

$$F_t(t, x; \tau) = \left(\frac{\partial f}{\partial t} + \sum_{i=1}^{d} \left(\frac{\partial f}{\partial \Pi_i}\right)\left(\frac{\partial \Pi_i}{\partial t}\right)\right)(t, x; \tau), \tag{9}$$

and

$$r_n(M, a) = \int_0^a \exp(Mu) u^n du \text{ for any } a{>}0 \text{ and square matrix } M.$$

For the numerical computation of this last integral see [2].

A naive approximation to $x(t - \tau_i)$ is the step function

$$\Pi_i(t) = x_{t_j}, \text{ for } (t - \tau_i) \in [t_j, t_{j+1}) \tag{10}$$

which is a practical approximation when the step-size Δ is very small. If it is not the case, a better approximation for $x(t - \tau_i)$ is the LL approximation (8) itself, that is

$$\begin{aligned} \Pi_i(t) &= x_{t_j} + r_0(F_x(t_j, x_{t_j}; \tau), \delta) F(t_j, x_{t_j}; \tau) \\ &\quad \left\{\delta r_0(F_x(t_j, x_{t_j}; \tau), \delta) - r_1(F_x(t_j, x_{t_j}; \tau), \delta)\right\} F_t(t_j, x_{t_j}; \tau) \end{aligned} \tag{11}$$

for $(t - \tau_i) \in [t_j, t_{j+1})$, where $\delta = t - \tau_i - t_j$. This last approximation became more attractive since the uniform convergence of the LL schemes for ODE have been proved in [15] (considering that the SDEs reduce to ODEs by setting the diffusion coefficient equal to zero). In this case, the order of the uniform convergency coincides with the global order of convergence of the LL scheme obtained in [14] for ODE. This is important since it has been proved that if an original ODE method is consistent of order p, it is enough to take an interpolation approximation of order $p-1$ to have a method of order p for DDE [3].

The numerical scheme (8),(11) has also the followings additional advantages: 1) they can be computed in the forward direction, that is, the value of $x(t-\tau_i)$ for any $(t-\tau_i) \in [t_k, t_{k+1})$ can be computed without using the value $x_{t_{k+1}}$; and 2) not more spurious information than those introduced by the approximation of the LL scheme itself is added in the approximation of $x(t-\tau_i)$.

When the approximation (10) is used then,

$$\frac{d\Pi_i}{dt}(t) = \left\{ \begin{array}{l} \frac{d\varphi}{dt}(t-\tau_i),\ for\ (t-\tau_i) \leq t_0 \\ 0,\ \text{otherwise} \end{array} \right\}$$

in the LL scheme (8). Instead we have

$$\frac{d\Pi_i}{dt}(t) = \left\{ \begin{array}{l} \frac{d\varphi}{dt}(t-\tau_i),\ for\ (t-\tau_i) \leq t_0 \\ \exp(F_x(t_j, x_{t_j}; \tau)\delta) F(t_j, x_{t_j}; \tau),\ for\ (t-\tau_i) \in [t_j, t_{j+1}) \end{array} \right\}$$

when the approximation (11) is used.

Other approaches that were considered to approximate the functions $x(t-\tau_i)$ are:

a) Defining the variable $z(t,\tau_i) = x(t-\tau_i)$ and add the equation $\frac{\partial z}{\partial t} = -\frac{\partial z}{\partial \tau_i}$ for each $i = 1, .., d$. Then, to approximate the right hand side of the first order partial differential equations by some difference scheme and apply the LL scheme to the resultant system of ODEs. Disadvantage: the dimension of the above system of ODEs can became too high.

b) To use the backward DDE to compute $x(t-\tau_i)$. Disadvantage: the backward DDEs has in general no solution or non-unique solution [4, 6].

c) To use the backward LL scheme for ODE to compute $x(t-\tau_i)$. Disadvantage: the resulting implicit scheme can has non-unique solution since in general the right hand side of a DDE is a non-injective function.

Comments:

The approximation (6)-(7) have been widely used for the numerical integration of DDEs, see e.g., [3, 7, 8, 9, 18, 19, 26]. They use several numerical schemes to solve the ODE (6) (e.g., Euler, Runge-Kutta, multi-steps methods) and several ways to approximate $x(t-\tau_i)$ in (7) (e.g., polynomial functions, θ-methods, continuous extensions of Runge-Kutta methods). The majority of them don't have the properties 1) and 2) of the numerical scheme (8),(11).

3 Local Linearization filters for delay state space models

Lets consider, first, the Ordinary State Space (OSS) model with continuous state equation

$$dx(t) = f(t, x(t))dt + \sigma(t)dW(t), \text{ for } t \geq t_0, \tag{12}$$

and discrete measurement equation

$$z_{t_k} = h(t_k, x(t_k)) + e_{t_k}. \tag{13}$$

where $e_{t_k} \sim N(0, \Sigma)$ i.i.d. vector of random variables for every $t_k = t_0 + k\Delta$, with $\Delta > 0$ and $k = 0, .., N$.

The classical extended Kalman filter for OSS model is given by the following theorem [10, pag. 278]:

Theorem 1: *The extended Kalman filter for the nonlinear system (12)-(13) consists in the prediction via*

$$x_{t_{k+1}/t_k} = x_{t_k/t_k} + \int\limits_{t_k}^{t_{k+1}} f(t, x_{t/t_k})dt, \tag{14}$$

$$P_{t_{k+1}/t_k} = \Phi(t_{k+1}, t_k, x_{t_k/t_k}) P'_{t_k/t_k} \Phi(t_{k+1}, t_k, x_{t_k/t_k}) + Q_{t_{k+1}}, \tag{15}$$

and at an observation

$$x_{t_{k+1}/t_{k+1}} = x_{t_{k+1}/t_k} + K_{t_{k+1}}(z_{t_{k+1}} - h(t_{k+1}, x_{t_{k+1}/t_k})), \tag{16}$$

$$P_{t_{k+1}/t_{k+1}} = P_{t_{k+1}/t_k} - K_{t_{k+1}} h_x(t_{k+1}, x_{t_{k+1}/t_k}) P_{t_{k+1}/t_k}. \tag{17}$$

Here the state transition matrix $\Phi(t_{k+1}, t_k, x_{t_k/t_k})$, the Kalman gain $K_{t_{k+1}}$ and the system noise covariance matrix $Q_{t_{k+1}}$ are given by the expressions

$$\Phi(t_{k+1}, t_k, x_{t_k/t_k}) = \exp(\int\limits_{t_k}^{t_{k+1}} f_x(u, x_{t_k/t_k})du), \tag{18}$$

$$K_{t_{k+1}} = P_{t_{k+1}/t_k} h'_x(t_{k+1}, x_{t_{k+1}/t_k})) \left(h_x(t_{k+1}, x_{t_{k+1}/t_k}) P_{t_{k+1}/t_k} h'_x(t_{k+1}, x_{t_{k+1}/t_k}) + \Sigma \right)^{-1}, \tag{19}$$

and

$$Q_{t_{k+1}} = \int_{t_k}^{t_{k+1}} \exp(f_x(t_k, x_{t_k/t_k})(u-t_k))\sigma(u)\sigma'(u)\exp(f_x(t_k, x_{t_k/t_k})(u-t_k))' du, \tag{20}$$

where the matrices functions h_x and f_x denote, respectively, the derivatives of the vector functions f and h with respect to x.

However, in general, this theorem can not be applied directly to the OSS model *(12)-(13)* due to expressions (14), (18) and (20) can only be explicitly computed in very few particular cases. Nevertheless, this theorem can be still useful to compute the filtered estimate of x in the general OSS model like *(12)-(13)* if the expressions (14), (18) and (20) are substituted by adequate approximations.

A class of such approximation have been proposed for the construction of the extended LL filtering algorithm considered in [11].

This extended LL filter is straightforward derived from the above theorem and the following three approximations: 1) the approximation to the solution of the ODE that describe the state prediction x_{t_{k+1}/t_k} in (14), that it,

$$dx_{t/t_k} = f(t, x_{t/t_k})dt, \;\; for\; t \in [t_k, t_{k+1}]$$

given by the expression

$$\begin{aligned} x_{t/t_k} \;&\approx\; x_{t_k/t_k} + r_0(f_x(x_{t_k/t_k}, t_k), t-t_k) f(x_{t_k/t_k}, t_k) \\ &\quad + \left\{(t-t_k) r_0(f_x(x_{t_k/t_k}, t_k), t-t_k) - r_1(f_x(x_{t_k/t_k}, t_k), t-t_k)\right\} f_t(x_{t_k/t_k}, t_k), \end{aligned}$$

which is obtained by setting the diffusion coefficient equal to zero in any one of the LL schemes for non-autonomous SDEs [2, 11, 24]; 2) the approximation to the system noise covariance matrix $Q_{t_{k+1}}$ given by the expression

$$vec(Q_{t_{k+1}}) \approx exp(M(x_{t_k/t_k}, t_n)\Delta) \sum_{i,j=0}^{m} r_{i+j}(-M(x_{t_k/t_k}, t_n), \Delta) vec(\Sigma_{i,j}(t_n)), \tag{21}$$

which is obtained by using a truncated Taylor expansion of the diffusion coefficient σ in (20) [13]. Here $M(x_{t_k/t_k}, t_n) = f_x(x_{t_k/t_k}, t_n) \oplus f_x(x_{t_k/t_k}, t_n)$, $\Sigma_{i,j}(t_n) = \frac{1}{(i!j!)}\sigma^{(i)}(t_n)(\sigma^{(j)})'(t_n)$ and $\sigma^{(i)}$ denotes the i-th derivatives of σ with respect to the variable t. The symbol $\oplus$ denotes the Kronecker sum and *vec* denotes the vec operator; 3) the approximation of the function $f_x(u, x_{t_k/t_k})$ in expression (18) by the constant $f_x(t_k, x_{t_k/t_k})$ for all $t \in [t_k, t_{k+1})$. Therefore, the state transition matrix is approximeted by the expression

$$\Phi(t_{k+1}, t_k, x_{t_k/t_k}) \approx \exp(f_x(t_k, x_{t_k/t_k})\Delta).$$

Thus, given the discrete observation set $\{z_{t_0}, .., z_{t_N}\}$ and the values x_{t_0/t_0} and P_{t_0/t_0}, the extended LL filter compute the filtered estimate x_{t_k/t_k} and its variance P_{t_k/t_k} for $k = 1, ..N$, by means of the following six-steps iterative algorithm:

1) one-step prediction

$$\begin{aligned} x_{t_{k+1}/t_k} &= x_{t_k/t_k} + r_0(f_x(x_{t_k/t_k}, t_k), \Delta) f(x_{t_k/t_k}, t_k) \\ &\quad + \left\{ \Delta r_0(f_x(x_{t_k/t_k}, t_k), \Delta) - r_1(f_x(x_{t_k/t_k}, t_k), \Delta) \right\} f_t(x_{t_k/t_k}, t_k), \end{aligned} \tag{22}$$

2) covariance of the above prediction

$$P_{t_{k+1}/t_k} = \exp(f_x(t_k, x_{t_k/t_k})\Delta) P_{t_k/t_k} \exp(f_x(t_k, x_{t_k/t_k})\Delta)' + Q_{t_{k+1}}, \tag{23}$$

where the matrix $Q_{t_{k+1}}$ is approximted by the expression (21),

3) innovation

$$\upsilon_{t_{k+1}} = z_{t_{k+1}} - h(t_{k+1}, x_{t_{k+1}/t_k}), \tag{24}$$

4) Kalman gain

$$K_{t_{k+1}} = P_{t_{k+1}/t_k} h'_x(t_{k+1}, x_{t_{k+1}/t_k}))(\Sigma^{\nu}_{t_{k+1}})^{-1}, \tag{25}$$

where the matrix

$$\Sigma^{\nu}_{t_{k+1}} = h_x(t_{k+1}, x_{t_{k+1}/t_k}) P_{t_{k+1}/t_k} h'_x(t_{k+1}, x_{t_{k+1}/t_k}) + \Sigma \tag{26}$$

is the innovation covariance matrix ,

5) filter estimate

$$x_{t_{k+1}/t_{k+1}} = x_{t_{k+1}/t_k} + K_{t_{k+1}} \upsilon_{t_{k+1}}, \tag{27}$$

6) covariance of the above filter estimate

$$P_{t_{k+1}/t_{k+1}} = P_{t_{k+1}/t_k} - K_{t_{k+1}} h_x(t_{k+1}, x_{t_{k+1}/t_k}) P_{t_{k+1}/t_k}. \tag{28}$$

This algorithm has been extensively used to filtering time series generated by highly nonlinear dynamical system such that Van der Pool and Two-Disc Dynamo in chaotic regimen of work [11, 23]. Also it has been shown to be highly effective in the estimation of non-observable series [25]. This algorithm is a generalization of the LL filter introduced early by Ozaki in 1993 [20] which has also been shown to be highly effective in filtering simulated and real time series in situations where other algorithms fails [20]-[23].

Lets us return to the DSS model with state equation

$$
\begin{aligned}
dx(t) &= f(t, x(t), x(t-\tau_1), .., x(t-\tau_d))dt + \sigma(t)dW(t), \text{ for } t \geq t_0, && (29)\\
x(t) &= \varphi(t), \text{ for -max}_{i=1..d}\{\tau_i\} \leq t \leq t_0 && (30)
\end{aligned}
$$

and discrete measurement equation

$$
z_{t_k} = h(t_k, x(t_k)) + e_{t_k}. \tag{31}
$$

If the state equation (29)-(30) is replaced by the equation

$$
dx(t) = f(t, x(t), \Pi_1(t), .., \Pi_d(t))dt + \sigma(t)dW(t), \text{ for } t \geq t_0, \tag{32}
$$

with

$$
\Pi_i(t) = \left\{ \begin{array}{l} \varphi(t), \ for \text{ -max}_{i=1..d}\{\tau_i\} \leq t \leq t_0 \\ an \ approximation \ to \ x(t-\tau_i), \ for \ t \geq t_0 \end{array} \right\}, \ for \ i = 1, .., d, \tag{33}
$$

then, the filtered estimate of x obtained by the LL filtering algorithm (22)-(28) for the OSS model (32),(31) is an approximation to the filtered estimate of x in the DSS model (29)-(31) . Note that in this case expression (22) takes the form of expression (8), and in the expressions (21) and (23) the function $f_x(t, x)$ should be replaced by $F_x(t, x; \tau)$ as it is defined in section 2.

Note that the theorem 1 can be directly applied to the OSS model (32),(31) if the approximation to $x(t - \tau_i)$ in (33) is non-random. However, if it is not the case, the theorem can still be applied since a random approximation for $x(t-\tau_i)$ does not affect the proof of the theorem (see [10, p. 272-278]). Moreover, a theorem similar to theorem 1 can be deduced if no approximation for $x(t - \tau_i)$ is made. In this case, in the expressions (14)-(20) the smoothed estimated $x_{t_k-\tau_i/t_k}$ is involved, for every $i = 1, .., d$. This last approach become more attractive since the smoothed estimated $x_{t_k-\tau_i/t_k}$ contents more information, and it has a better theoretical support.

4 Identification of delay state space models

Lets us consider, first, the identification of the OSS model with continuous state equation

$$
dx(t) = f(t, x(t), \theta)dt + \sigma(t, \theta)dW(t), \text{ for } t \geq t_0, \tag{34}
$$

and discrete measurement equation

$$z_{t_k} = h(t_k, x(t_k), \theta) + e_{t_k}. \tag{35}$$

where $e_{t_k} \sim N(0, \Sigma(\theta))$ i.i.d. vector of random variables for every $t_k = t_0 + k\Delta$, with $\Delta > 0$ and $k = 0, .., N$.

The identification problem here is the following: *given the above state space model and a set of discrete observations* $\{z_{t_0}, .., z_{t_N}\}$, *to estimate both the non-observable components of the state vector* x *at the instant of time* $t_0, .., t_N$ *and the unknown set of parameters* θ.

An effective method to solve this identification problem has been proposed by Ozaki [21]. It is based on the maximization of the innovation form of the likelihood function for OSS models and it involves the following two steps:

1. The computation of the estimates $\widehat{\theta}$, $\widehat{x}_{t_0/t_0}$ and $\widehat{P}_{t_0/t_0}$ by minimizing the negative log likelihood

$$-2\ln(L(z; \theta, x_{t_0/t_0}, P_{t_0/t_0})) = -\ln(2\pi)N + \sum_{k=0}^{N} \left\{ \ln \left| \Sigma^{\nu}_{t_{k+1}} \right| + \upsilon'_{t_{k+1}} (\Sigma^{\nu}_{t_{k+1}})^{-1} \upsilon_{t_{k+1}} \right\} \tag{36}$$

   with respect to the set of unknown variables $\theta, x_{t_0/t_0}$ and P_{t_0/t_0}. Here $\upsilon_{t_{k+1}}$ and $\Sigma^{\nu}_{t_{k+1}}$ are, respectively, the innovation vector (24) and its covariance matrix (26) defined in the LL filtering algorithm (22)-(28).

2. The computation of the filtered estimate of x by means of the LL filtering algorithm (22)-(28) applied to the OSS model (34)-(35) with $\theta = \widehat{\theta}$, $x_{t_0/t_0} = \widehat{x}_{t_0/t_0}$ and $P_{t_0/t_0} = \widehat{P}_{t_0/t_0}$.

This method has been extensively used with effectiveness in the model identification from a time series generated by highly nonlinear dynamical system such as Van der Pool and Two-Disc Dynamo in chaotic regimen of work [11, 23]. Also it has been shown to be highly effective in the identification of a high dimensional neural mass model from real EEG time series [25] and in the seasonal adjustement of time series of official statistics [22]. Now, lets us return to main subject of this report, the identification of the delay state space model (1)-(3). If the state equation (1)-(2) is replaced by the equation

$$dx(t) = f(t, x(t), \Pi_1(t), .., \Pi_d(t), \theta)dt + \sigma(t, \theta)dW(t), \text{ for } t \geq t_0, \tag{37}$$

with

$$\Pi_i(t) = \left\{ \begin{array}{l} \varphi(t), \ \mathit{for} \ \text{-}\max_{i=1..d}\{\tau_i\} \leq t \leq t_0 \\ \mathit{an\ approximation\ to}\ x(t-\tau_i), \ \mathit{for}\ t \geq t_0 \end{array} \right\}, \ \mathit{for}\ i = 1, .., d, \tag{38}$$

then, Ozaki's identification method described above can be applied to the OSS model (37),(3). However, a new problem arise in this case, that is, the function φ in the expression (38) is also unknown and should be estimated as well as x_{t_0/t_0} and P_{t_0/t_0}. Nevertheless, Ozaki's identification procedure can still be applied to this case with some minor modifications.

Suppose that function φ can be approximated by a parametric function $\Phi(\vartheta)$ in the interval $[t_0 - T, t_0]$, where T is chosen in such a way that the relation $t_0 - \tau_i \in [t_0 - T, t_0]$ holds for the range of all possible values of τ_i and for $i = 1, .., d$. Then, the identification procedure for DSS models has the following steps:

1. The computation of the estimates $\widehat{\tau}$, $\widehat{\theta}$, $\widehat{\varphi}$ and $\widehat{P}_{t_0/t_0}$ by minimizing the negative log likelihood

$$-2\ln(L(z;\tau,\theta,\vartheta,P_{t_0/t_0})) = -\ln(2\pi)N + \sum_{k=0}^{N}\left\{\ln\left|\Sigma^{\nu}_{t_{k+1}}\right| + \upsilon'_{t_{k+1}}(\Sigma^{\nu}_{t_{k+1}})^{-1}\upsilon_{t_{k+1}}\right\} \tag{39}$$

   with respect to the set of unknown variables τ, θ, φ and P_{t_0/t_0}. Here $\upsilon_{t_{k+1}}$ and $\Sigma^{\nu}_{t_{k+1}}$ are, respectively, the innovation vector (24) and its covariance matrix (26), and $x_{t_0/t_0} = \Phi(t_0, \vartheta)$.

2. The computation of the filtered estimate of x by means of the LL filtering algorithm (22)-(28) applied to the OSS model (37),(3) with $\tau = \widehat{\tau}$, $\theta = \widehat{\theta}$, $\varphi = \widehat{\varphi}$, $P_{t_0/t_0} = \widehat{P}_{t_0/t_0}$, $x_{t_0/t_0} = \Phi(t_0, \widehat{\vartheta})$ and $\varphi = \Phi(\widehat{\vartheta})$.

In the above procedure there is no restriction to the approximation of the function φ. Thus, the class of approximating function Φ can be chosen for each particular DSS model, which is a logical advantage. However, to construct the approximating function Φ no information on the structure of the state equation is used. Therefore, the performance of identification procedure strongly depends on the performance of the approximation of φ by $\Phi(\vartheta)$ in the interval $[t_0 - T, t_0]$.

One way to reduce the effect of a possible unsatisfactory approximation of φ in the interval $[t_0 - T, t_0]$ is based on the generation of a *reference (or nominal)* deterministic trajectory of the state equation (1)-(2) . Such a method supposes that the state equation (1)-(2) are well defined for every $t \geq (t_a - T)$ where $t_a \ll (t_0 - T)$, that is,

$$dx(t) = f(t, x(t), x(t-\tau_1), .., x(t-\tau_d), \theta)dt + \sigma(t,\theta)dW(t), \text{ for } t \geq t_a, \tag{40}$$

$$x(t) = \varphi_a(t), \text{ for } (t_a - T) \leq t \leq t_a. \tag{41}$$

Then, the *reference* deterministic trajectory $\overline{x}$ is obtained from the LL numerical integration of differential equation (40)-(41) with $\sigma(t,\theta) = 0$ and $\varphi_a = \Phi(\vartheta)$, where Φ is a parametric function and ϑ is a set of parameters. Finally, the function φ is approximated by $\overline{x}(\vartheta)$ in the interval $[t_0 - T, t_0]$.

In this case the identification procedure for the DSS model has the following steps:

1. The computation of the estimates $\widehat{\tau}$, $\widehat{\theta}$, $\widehat{\varphi}$ and $\widehat{P}_{t_0/t_0}$ by minimizing the negative log likelihood

$$-2\ln(L(z;\tau,\theta,\vartheta,P_{t_0/t_0})) = -\ln(2\pi)N + \sum_{k=0}^{N}\left\{\ln\left|\Sigma^{\nu}_{t_{k+1}}\right| + \upsilon'_{t_{k+1}}(\Sigma^{\nu}_{t_{k+1}})^{-1}\upsilon_{t_{k+1}}\right\}$$

   with respect to the set of unknown variables τ, θ, φ and P_{t_0/t_0}. Here $\upsilon_{t_{k+1}}$ and $\Sigma^{\nu}_{t_{k+1}}$ are, respectively, the innovation vector (24) and its covariance matrix (26), and $x_{t_0/t_0} = \overline{x}(t_0,\vartheta)$.

2. The computation of the filtered estimate of x by means of the LL filtering algorithm (22)-(28) applied to the OSS model (37),(3) with $\tau = \widehat{\tau}$, $\theta = \widehat{\theta}$, $\varphi = \widehat{\varphi}$, $P_{t_0/t_0} = \widehat{P}_{t_0/t_0}$, $x_{t_0/t_0} = \overline{x}(t_0,\widehat{\vartheta})$ and $\varphi = \overline{x}(\widehat{\vartheta})$.

Other approaches that were considered to approximate the functions φ are:

a) The smoothness approach. It is based in the use of the smoothed estimate of x as an approximation to the values of the unknown function φ at the points $t_k = t_0 + k\Delta$ for $k = -1,-2,..,-M$, where t_{-M} is chosen in such a way that the relation $t_0 - \tau_i \in [t_{-M}, t_0]$ holds for the range of all possible values of τ_i and for $i = 1,..,d$. In this case, the identification procedure is as follows. Suppose that, initially, the values z_{t_k}, x_{t_{k+1}/t_k} and P_{t_{k+1}/t_k} are given for $k = -M,..,-1$. Then, for each new value of x_{t_0/t_0} and P_{t_0/t_0} the values of function φ at the points t_k are computed in four steps:

1. Compute the values x_{t_k/t_k} and P_{t_k/t_k} for $k = -M,..,-1$ by using the expression (24)-(28)

2. Compute the smoothed estimate x_{t_k/t_0} and P_{t_k/t_0} for $k = -M,..,-1$ by using the smoothing algorithm [16]:

$$x_{t_k/t_0} = x_{t_k/t_k} + J_{t_k/t_k}(x_{t_{k+1}/t_0} - x_{t_{k+1}/t_k}),$$

$$P_{t_k/t_0} = P_{t_k/t_k} + J_{t_k/t_k}(P_{t_{k+1}/t_0} - P_{t_{k+1}/t_k})J'_{t_k/t_k},$$

$$J_{t_k/t_k} = P_{t_k/t_k}\exp(F_x(t_k, x_{t_{k+1}/t_{k+1}};\tau,\theta)\Delta)' P^{-1}_{t_{k+1}/t_k},$$

3. Update $\varphi(t_k)$ with the estimate x_{t_k/t_0} for $k = -M, .., -1$

4. Update x_{t_{k+1}/t_k} and P_{t_{k+1}/t_k} with the estimates x_{t_k/t_0} and P_{t_k/t_0} for $k = -M, .., -1$

Disadvantage: $F_x(t_k, x_{t_{k+1}/t_{k+1}}; \tau, \theta)$ can not be computed.

b) To use the backward DDE. Disadvantage: the same as that in section 2.

c) To use the backward LL scheme for ODE. Disadvantage: the same as that in section 2.

5 Local Linearization schemes for the numerical integration of delay stochastic differential equations

Lets us consider the delayed stochastic differential equation

$$\begin{aligned} dx(t) &= f(t, x(t), x(t-\tau_1), .., x(t-\tau_d))dt + \sigma(t)dW(t),\ t \geq t_0, & (42) \\ x(t) &= \varphi(t),\ for\ \text{-max}_{i=1..d}\{\tau_i\} \leq t \leq t_0 & (43) \end{aligned}$$

where $x(t) \in R^n$, $\sigma(t) \in R^{n \times m}$, and W is an m-dimensional Wiener process. The above equation can be approximated by the following non-autonomous Stochastic Differential Equation (SDE)

$$dx(t) = f(t, x(t), \Pi_1(t), .., \Pi_d(t))dt + \sigma(t)dW(t),\ \text{for } t \geq t_0,$$

where

$$\Pi_i(t) = \left\{ \begin{array}{l} \varphi(t),\ for\ \text{-max}_{i=1..d}\{\tau_i\} \leq t \leq t_0 \\ an\ approximation\ to\ x(t-\tau_i),\ for\ t \geq t_0 \end{array} \right\} \tag{44}$$

for $i = 1, .., d$.

Therefore, a LL scheme for non-autonomous SDEs (e.g., [2, 12, 24]) can be used for the numerical integration of that SDE and in this way obtained a numerical scheme for the delay SDE (42)-(43). Lets us consider, for example, the LL scheme [12]:

$$\begin{aligned} x_{t_{k+1}} = {} & x_{t_k} + r_0(F_x(t_k, x_{t_k}; \tau), \Delta)F(t_k, x_{t_k}; \tau) \\ & + \{\Delta r_0(F_x(t_k, x_{t_k}; \tau), \Delta) - r_1(F_x(t_k, x_{t_k}; \tau), \Delta)\} \\ & \times \left(F_t(t_k, x_{t_k}; \tau) + \frac{1}{2} \sum_{i,j=1}^{n} [\sigma(t_k)\sigma'(t_k)]^{i,j}\ F_{xx}^{i,j}(t_k, x_{t_k}; \tau) \right) \\ & + \xi(t_k, t_{k+1}) \end{aligned} \tag{45}$$

where the functions r_0, r_1, F, F_x and F_t are defined as in section 2,

$$F_{xx}^{i,j}(t,x;\tau) = \frac{\partial^2 f}{\partial x^i \partial x^j}(t,x,\Pi_1(t),..,\Pi_d(t))$$

and

$$\begin{aligned}
\xi(t_k,t_{k+1}) &= \int\limits_{t_k}^{t_{k+1}} \exp(F_x(t_k,x_{t_k};\tau)(t_{k+1}-u))\sigma(u)dW(u) \\
&= \sigma(t_{k+1})W(t_{k+1}) - \exp(F_x(t_k,x_{t_k};\tau)\Delta)\sigma(t_k)W(t_k) \\
&\quad - \sum_{i=0}^{R-1} \int\limits_{h_i}^{h_{i+1}} \left(\frac{\partial \exp(F_x(t_k,x_{t_k};\tau)(t_{k+1}-u))\sigma(u)}{\partial u} \right) W(u)du,
\end{aligned}$$

with $h_i = t_k + i\frac{\Delta}{R}$, *and* $R \sim \frac{1}{\Delta}$. Here, the last integral is approximated by means of the Trapezoidal rule. A naive approximation to $x(t-\tau_i)$ in (44) is the step function

$$\Pi_i(t) = x_{t_j}, \text{ for } (t-\tau_i) \in [t_j, t_{j+1}) \tag{46}$$

which is a practical approximation when the step-size Δ is very small. However, a better approximation to $x(t-\tau_i)$ similar to the expression (11), obtained by considering the LL approximation itself, seems to be impossible since the process ξ in LL scheme for SDEs is not derivable. Nevertheless, the LL approximation can still to be useful to construct a random approximation to $x(t-\tau_i)$ for $(t-\tau_i) \in [t_j, t_{j+1})$ and $j = 0,1,..,N$.

Let $S^k = \{(t_l - \tau_i - t_k) : (t_l - \tau_i) \in [t_k - t_{k+1})$, for $l = K+1,..,N$ and $i = 1,..,d\} \cup \{0,\Delta\}$ be a set of numbers, δ^k be a sequence of numbers formed by the elements of S^k ordered in ascending way, and δ_j^k be the j-th element of δ^k.

In a scheme with fixed step-size, like (45), the set of points $\{(t_l-\tau_i)\}_{l=0,..,N}^{i=1,..,d}$ at which the function x is going to be evaluated is priori known. Moreover, the set of points at which the Wienner process W should be evaluated to compute $\xi(t_k, t_k+\delta_j^k)$, for every k and j, is also priori known. Therefore, an adequate sequence of the values of the Wienner process can be priori constructed, and in such a way, the following approximation can be constructed

$$\Pi_i(t) = x_{t_k+\delta_j^k}, \text{ for } (t-\tau_i) \in [t_k+\delta_j^k, t_k+\delta_{j+1}^k) \tag{47}$$

where

$$x_{t_k+\delta_j^k} = x_{t_k} + r_0(F_x(t_k,x_{t_k};\tau),\delta_j^k)F(t_k,x_{t_k};\tau)$$

$$\left\{\delta_j^k r_0(F_x(t_k, x_{t_k}; \tau), \delta_j^k) - r_1(F_x(t_k, x_{t_k}; \tau), \delta_j^k)\right\}$$
$$\times\left(F_t(t_k, x_{t_k}; \tau) + \frac{1}{2}\sum_{i,j=1}^{n}[\sigma(t_k)\sigma'(t_k)]^{i,j}\ F_{xx}^{i,j}(t_k, x_{t_k}; \tau)\right)$$
$$+\xi(t_k, t_k + \delta_j^k).$$

This last approximation became more attractive since the mean-square uniform convergence of the LL scheme (45) have been proved in [15]. In both case, when the approximations (46) or (47) are used then,

$$\frac{d\Pi_i}{dt}(t) = \left\{\begin{array}{l} \frac{d\varphi}{dt}(t-\tau_i),\ for\ (t-\tau_i) \leq t_0 \\ 0,\ \text{otherwise} \end{array}\right\}.$$

Therefore,

$$F_t(t, x; \tau) = \frac{\partial f}{\partial t}(t, x; \tau),$$

in the LL scheme (45). Other approaches that were considered to approximate the functions $x(t-\tau_i)$ are:

a) Defining the variable $z(t, \tau_i) = x(t-\tau_i)$ and add the equation $d_t z = -d_{\tau_i} z$ for each $i = 1, .., d$. Then, to approximate the right hand side of the stochastic evolution equations by some difference scheme and apply the LL scheme to the resultant system of SDEs. Disadvantage: the dimension of the above system of SDEs can became too high.

6 Numerical Experiments

In this section some numerical experiments are presented. In particular, the performance of the simplest identification method is partially evaluated. That is, the approximation (10) is used in the extended LL filtering algorithm, and the expression (39) is used to estimate the unknown parameters. Therefore, the approximation (46) is used in the LL scheme (45) to generate simulated time series for the DSS model that is going to be considered.

Let us consider the identification of the 11−dimensional DSS model with continuous state equation

$$\begin{aligned} dx_1(t) &= (\alpha_e x_1(t) + \beta_e x_5(t) + \varepsilon_e c_4 g(x_6(t-\tau)) + c_2\alpha x_4(t) \\ &\quad + c_2\beta x_8(t) + c_2\varepsilon_i g(x_7(t)) + \varepsilon_e\pi)dt + \sigma dW(t), \end{aligned} \tag{48}$$

$$dx_2(t) = (\alpha_e x_2(t) + \beta_e x_6(t) + \varepsilon_e c_3 g(x_5(t)))dt, \tag{49}$$

$$dx_3(t) = (\alpha_e x_3(t) + \beta_e x_7(t) + \varepsilon_e c_1 g(x_5(t)))dt, \tag{50}$$

$$dx_4(t) = (\alpha_i x_4(t) + \beta_i x_8(t) + \varepsilon_i g(x_7(t)))dt, \tag{51}$$

$$dx_5(t) = x_1(t)dt, \tag{52}$$

$$dx_6(t) = x_2(t)dt, \tag{53}$$

$$dx_7(t) = x_3(t)dt, \tag{54}$$

$$dx_8(t) = x_4(t)dt, \tag{55}$$

$$dx_9(t) = (k_2 x_9(t) + k_1 x_{11}(t) + k_0 x_{10}(t) + A\varpi_n^2 x_2(t))dt, \tag{56}$$

$$dx_{10}(t) = x_{11}(t)dt, \tag{57}$$

$$dx_{11}(t) = x_9(t)dt, \tag{58}$$

for $t \geq t_0$, and

$$x(t) = \varphi(t), \text{ for } (t_0 - \tau) \leq t \leq t_0. \tag{59}$$

Here, σ is a 11×11 matrix defined by

$$\sigma = \begin{bmatrix} \varepsilon^2 \sigma_p^2 & 0 \\ 0 & 0 \end{bmatrix}.$$

The discrete measurement equation is given by the expression

$$z_{t_k} = x_{10}(t_k) + e_{t_k}. \tag{60}$$

where $e_{t_k} \sim N(0, 0.001)$ for every $t_k = t_0 + k\Delta$, with $\Delta = 5 \times 10^{-4}$ sec, $k = 0, .., N$, and $N = 2560$. The state space model (48)-(60) formally represents a neural mass model that describes the EEG background activity of the brain plus a Butterworth filter of the recording system. For more details see [27] and [25].

The identification problem here is the following: *given the above delay state space model and a set of discrete observations* $\{z_{t_0}, .., z_{t_N}\}$, *to estimate both the non-observable components of the state vector* x *at the instant of time* $t_0, .., t_N$ *and the unknown set of parameters* c_1, c_2, c_3, c_4, π, σ_p^2, A and τ. (The value of the remained parameters are assumed to be the same as in [25]).

In order to evaluate the performance of the extended LL filtering algorithm lets us consider the parameters c_1, c_2, c_3, c_4, π, σ_p^2, A and τ given. In particular the value $c_1 = 10$, $c_2 = 1.5$, $c_3 = 36$, $c_4 = 7$, $\pi = 179.7$, $\sigma_p^2 = 15.84$ and $A = 2.8 \times 10^{-4}$ are set in the DSS model (48)-(60). Here, it is also assumed that function φ is known in $[t_0 - \tau, t_0]$. Specifically, φ is generated in $[t_0 - \tau, t_0]$ by the LL scheme (45) applied to the state equations (48)-(59) starting from $t \ll t_0 - \tau$ with $\varphi = 0$ in $[t - \tau, t]$.

Figure 1 shows both a trajectory of the state vector x generated by the LL scheme (45) and the filtered estimate of x obtained for the extended LL filtering algorithm for the DSS model (48)-(60), with $\tau = 50\Delta$. Note that the non-observed variables of x are well estimated by the filtering algorithm. In this case the estimated innovation process looks like a Gaussian white noise process with mean near to zero and a variance less than 1% of the total signal power (see Figure 2).

Similar results are obtained when the extended LL filtering algorithm is applied to the DSS model (48)-(60), with $\tau = 10\Delta$ and $\tau = 0$. See Figures 3 and 4, respectively. Figures 5a and 5b show the estimated innovation process obtained in each case.

Figure 6a shows the spectrum of the observed variable x_{10} of the Figure 1 and the spectrum a real EEG time series taking as reference. Figures 6b and 6c show the spectrum of the observed variable x_{10} of the Figure 3 and 4, respectively, and the spectrum a same real EEG time series taking as reference. Note that the simulated time series x_{10} generated by the LL scheme (45) has a spectrum similar to those of a real EEG time series and, in addition, the amplitude of the main peak of the spectrum of x_{10} decrease when the value of the delay parameter τ increase which, is in correspondence with the relation predicted in [27].

Now, we suppose that the parameters $c_1, c_2, c_3, c_4, \pi, \sigma_p^2, A$ and τ are unknown.

Table 1

$c_1 = 10$	$c_1^0 = 10.1$	$\widehat{c}_1 = 9.82$
$c_2 = 1.5$	$c_2^0 = 1.4$	$\widehat{c}_2 = 1.46$
$c_3 = 36$	$c_3^0 = 36.6$	$\widehat{c}_3 = 35.35$
$c_4 = 7$	$c_4^0 = 6.8$	$\widehat{c}_4 = 6.96$
$\pi = 179.7$	$\pi^0 = 180.0$	$\widehat{\pi} = 175.7$
$A = 2.8 \times 10^{-4}$	$A^0 = 3 \times 10^{-4}$	$\widehat{A} = 2.7 \times 10^{-4}$
$\sigma_p^2 = 15.8$	$\sigma_p^2 = 15.0$	$\widehat{\sigma_p^2} = 15.25$
$\tau = 0.0125$	$\tau^0 = 0.01$	$\widehat{\tau} = 0.0124$

Table I shows the estimated values $\widehat{c}_1$, $\widehat{c}_2$, $\widehat{c}_3$, $\widehat{c}_4$, $\widehat{\pi}$, $\widehat{\sigma}_p^2$, $\widehat{A}$ and $\widehat{\tau}$ for c_1, c_2, c_3, c_4, π, σ_p^2, A and τ, respectively obtained by minimizing expression (39) with $x_{t_0/t_0} = \varphi(t_0)$ and $P_{t_0/t_0} = 0$. The filtered estimate of x for the DSS model (48)-(60) with $c_1 = \widehat{c}_1$, $c_2 = \widehat{c}_2$, $c_3 = \widehat{c}_3$, $c_4 = \widehat{c}_4$, $\pi = \widehat{\pi}$, $\sigma_p^2 = \widehat{\sigma}_p^2$, $A = \widehat{A}$ and $\tau = \widehat{\tau}$ is as good as the filtered estimate shown in Figures 1,3 and 4. In this case the innovation process is similar to those obtained in Figures 2 and 5. See Appendix A for practical details about the parameter estimation.

Acknowledgment

The authors are grateful to the Ministry of Education, Science and Culture of Japan for supporting the visit of J.C. Jimenez to the Institute of Statistical Mathematics, Tokyo, during which the methods for the identification of DSS models presented in this report were introduced and developed.

7 Appendix A

The goal of this appendix is to provide the reader with a effective way to reproduce our results for the identification of DSS models since the minimization of the negative log likelihood (39), as any other objective function, is not easy task. The estimation recipe that will be explained here assumes that a local optimization procedure will be used. It is reasonable since in the majority of the practical identification problems there exists information about the possible values of the unknown parameters. The reader can omit this appendix if he has some general and efficient subroutine for computing the global minimum of a function.

The estimation recipe has the following steps:

1. Selecting some adequate initial values τ_0, θ_0 and ϑ_0 for the set of parameters τ, θ and ϑ, respectively. τ_0 and θ_0 are chosen from the set of possible parameter values for the particular DSS model under consideration. ϑ_0 is choose after that the state equation with $\tau = \tau_0$ and $\theta = \theta_0$ is numerically integrated in a large interval of time $[t_0 - T, t_0]$ starting with some initial estimate for φ , e.g., $\varphi(t) = \varphi_0(t)$, for $t \in [t_0 - T - \max_{i=1,..,d} \{(\tau_0)_i\}, t_0 - T]$. Then, θ_0 is estimated from computed values of x at every points $t_k \in [t_0 - \xi, t_0]$, where ξ is such that the relation $t_0 - \tau_i \in [t_0 - \xi, t_0]$ holds for the range of all possible values of τ_i and for $i = 1, .., d$.

2. Computing the extended LL filtering algorithm for the DSS model with $\tau = \tau_0$, $\theta = \theta_0$, $\vartheta = \vartheta_0$, $x_{t_0/t_0} = \Phi(t_0, \theta_0)$ and $P_{t_0/t_0} = c_0 I$. Here $c_0 \in [10^{-9}, 10^{-3}]$ and is I is the identity matrix.

3. Ploting the innovation function,

4. Selecting an instant of time t_m after that the transient effects that risen in the graph of the innovation function disappear,

5. Computing the estimates $\widehat{\tau}$ and $\widehat{\theta}$ by minimizing the expression

$$-2\ln(L(z;\tau,\theta/\vartheta_0,P_{t_0/t_0})) = -\ln(2\pi)N + \sum_{k=m}^{N} \left\{ \ln\left|\Sigma^{\nu}_{t_{k+1}}\right| + \upsilon'_{t_{k+1}}(\Sigma^{\nu}_{t_{k+1}})^{-1}\upsilon_{t_{k+1}} \right\}$$

with respect to the set of parameters τ and θ. The values τ_0 and θ_0 are used to initialize the local optimization subroutine.

6. Computing the estimate $\widehat{\vartheta}$ by minimizing the expression

$$-2\ln(L(z;\vartheta/\widehat{\tau},\widehat{\theta},P_{t_0/t_0})) = -\ln(2\pi)N + \sum_{k=1}^{N}\left\{\ln\left|\Sigma^{\nu}_{t_{k+1}}\right| + \upsilon'_{t_{k+1}}(\Sigma^{\nu}_{t_{k+1}})^{-1}\upsilon_{t_{k+1}}\right\}$$

with respect to the set of parameter ϑ. The value ϑ_0 used to initialize the local optimization subroutine. Here $P_{t_0/t_0} = c_1 I$, where $c_1 \ll c_0$.

The constant c_0 in the step (2) is selected by trial and error in such a way that the innovation function computed by the LL filtering algorithm has a large transient in time.

The minimization of the functions in the step (5) and (6) is carried out under the constraint that the mean and variance of the sequence $\{\upsilon_{t_1},..,\upsilon_{t_N}\}$ must be smaller than mean and variance of the sequence $\{\upsilon_{t_m},..,\upsilon_{t_N}\}$ computed in step (2).

We suggest the reader the excellent book [5, Chap. 8] for all the remaining important details concerning to the practical optimization.

References

[1] Baker T.H. and Paul C.A.H. (1994) Computing stability regions - Runge-Kutta methods for delay differential equations. IMA J. Numer. Anal., 14, 347-362.

[2] Biscay R., Jimenez J.C., Riera J.J., Valdes P.A. (1996) Local linearization method for the numerical solution of stochastic differential equations, Annals Inst. Statist. Math., 48, 631-644.

[3] Calvo M. and Grande T. (1988) On the asymptotic stability of θ-methods for delay differential equations. Numer. Math., 54, 257-269.

[4] Diekmann O., van Gils S.A., Verduyn Lunel S.M. and Walther H.O. (1995) Delay differential equations. Springer-Verlag.

[5] Gill P.E., Murray W. and Wright M.H. (1981) Practical Optimization. Academic Press.

[6] Hale J.K. and Verduyn Lunel S.M. (1993) Introduction to functional differential equations. Springer-Verlag.

[7] Higham D.J. and Famelis I.T. (1995) Equilibrium states of adaptive algorithms for delay differential equations. J Comput. Appl. Math., 58, 151-169.

[8] Hongjiong T. and Jiaoxun K. (1995) The stability of the θ-methods in the numerical solution of delay differential equations with several delay terms. J. Comput. Appl. Math., 58, 171-181.

[9] In't Hout K.J. (1992) A new interpolation procedure for adapting Runge-Kutta methods to delay differential equations. Bit, 32, 634-649.

[10] Jazwinski A.H. (1970) Stochastic processes and filtering theory. Academic Press.

[11] Jimenez J.C. (1996) Local Linealization method for the numerical solution of stochastic differential equations and nonlinear Kalman filters. In Proceedings of the Symposium: Studies on data analysis by statistical software, Fuji Kenshu-Jyo, Japan, Nov. 1995, 291-292.

[12] Jimenez J.C., Shoji I. and Ozaki T. (1996) Simulation of stochastic differential equations through the local linearization method. A comparative study. Submitted to J. Statist. Physics.

[13] Jimenez J.C., Valdes P.A., Rodriguez L.M., Riera J. and Biscay R. (1997) Computing the noise covariance matrix of the Local Linealization scheme for the numerical solution of SDEs. Appl. Math. Letters. (in press)

[14] Jimenez J.C., Mora C. and Biscay R. (1997) Properties of a Local Linearization method for the numerical integration of ordinary differential equations. Submitted to Math. Comput.

[15] Jimenez J.C. and Biscay R. (1997) Rate of convergence of some Local Linearization schemes for stochastic differential equations. In preparation.

[16] Meditch J.S. (1967) Orthogonal projection and discrete optimal linear smoothing. J. SIAM Control, 5, 74-89.

[17] Mohammed S.E.A (1984) Stochastic functional differential equations. Pitman Advanced Publishing Program.

[18] Neves K.W. (1975) Automatic integration of functional differential equations. ACM Trans. Math. Software, 1, 357-368.

[19] Orbele H.J. and Pesch H.J. (1981) Numerical treatment of delay differential equations by Hermite interpolation. Numer. Math., 37, 235-255.

[20] Ozaki T. (1993) A local linearization approach to nonlinear to nonlinear filtering. Int. J. Control, 57, 75-96.

[21] Ozaki T. (1994) The local linearization filter with application to nonlinear system identification. In Bozdogan H. (eds.) Proceedings of the First US/Japan Conference on Frontiers of Statistical Modeling: An informational approach. Kluwer Academic Publishers, 217-240.

[22] Ozaki T. and Thomson P.J. (1994) A dynamical system approach to X-11 type seasonal adjustment.. Ree. Memo. No. 498, The Institute of Statistical Mathematics, Tokyo, Japan.

[23] Ozaki T., Jimenez J.C. and Haggan-Ozaki V. (1997) The role of the likelihood function in the estimation of chaos models. Submitted to J. Time Series Anal.

[24] Shoji I. (1995) A local linearization method for multivariate stochastic processes, Res. Memo. No. 586, The Institute of Statistical Mathematics, Tokyo, Japan.

[25] Valdes P.A., Jimenez J.C., Riera J., Biscay R. and Ozaki T. (1997) Nonlinear EEG analysis based on a neural mass model. Submitted to Biol. Cyb.

[26] Zennaro M. (1986) Natural continuous extensions of Runge-Kutta methods. Math. Comput., 46, 119-133.

[27] Zetterberg L.H., Kristiansson L. and Mossberg (1978) Performance of a model for a local neuron population. Biol Cybernetics, 31, 15-26.

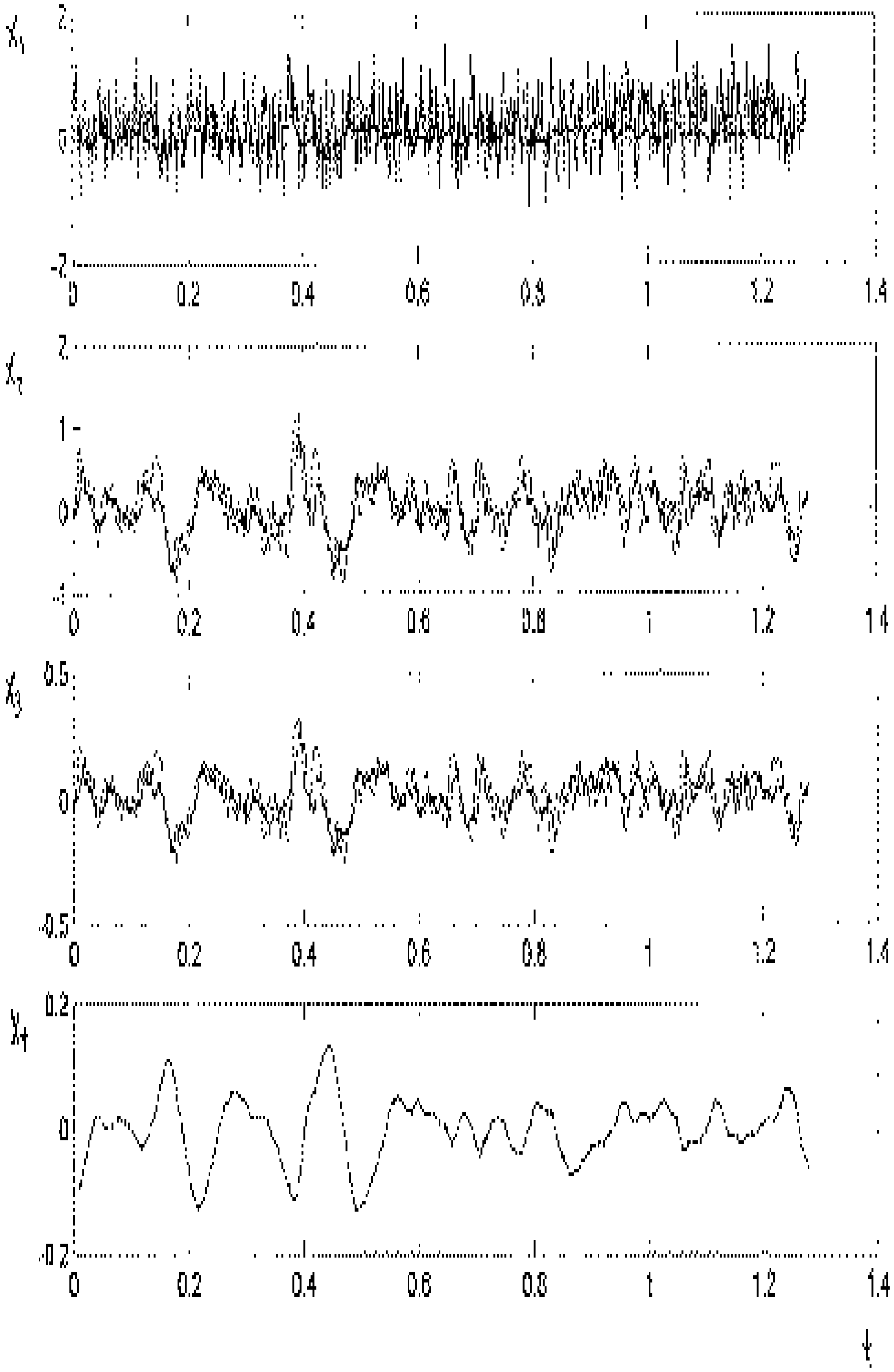


Figure 1a

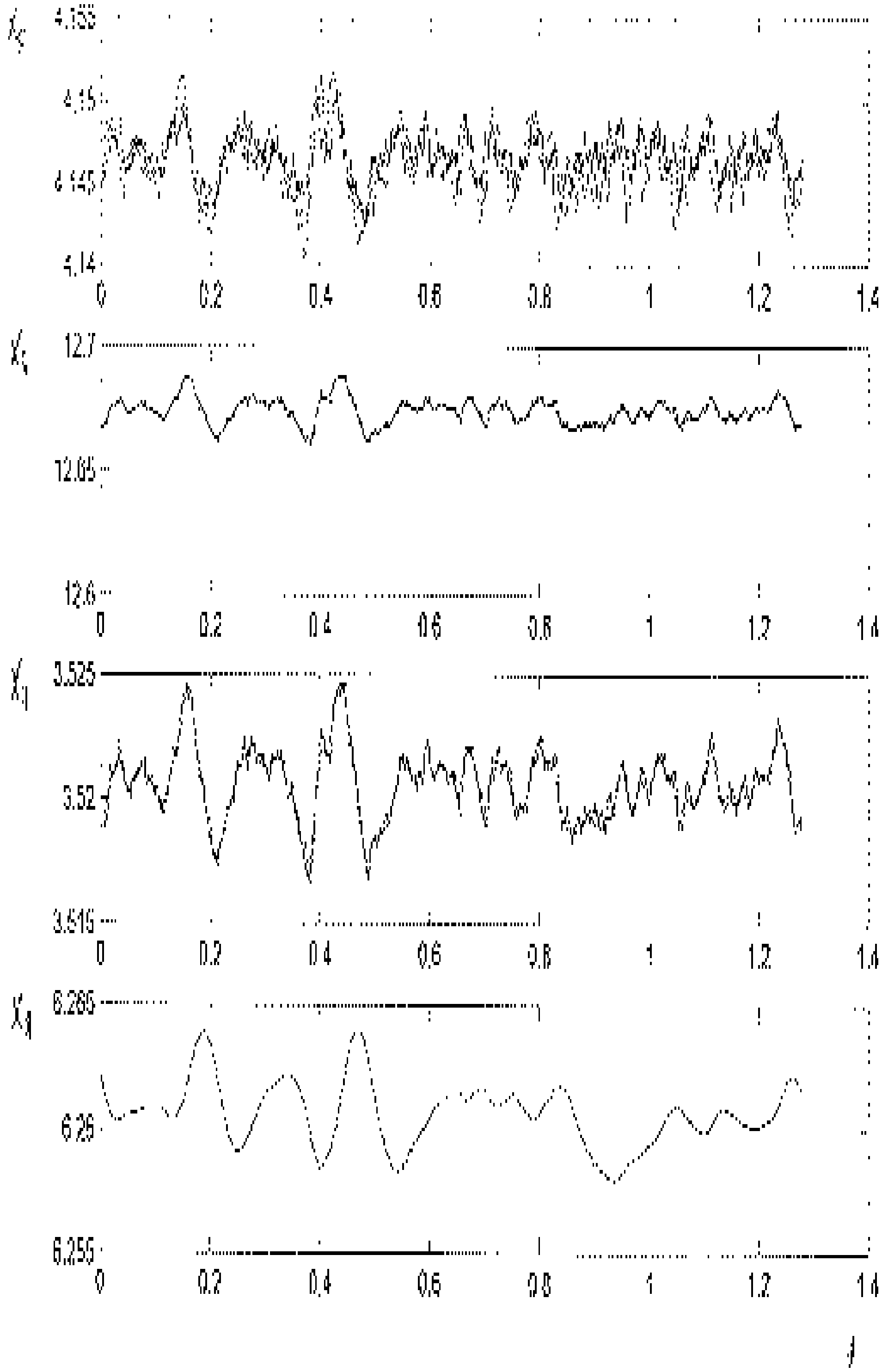

4.155
4.15
4.145
4.14
12.7
12.65
12.6
3.525
3.52
3.515
6.265
6.26
6.255
0
0.2
0.4
0.6
0.8
1
1.2
1.4

Figure 1b

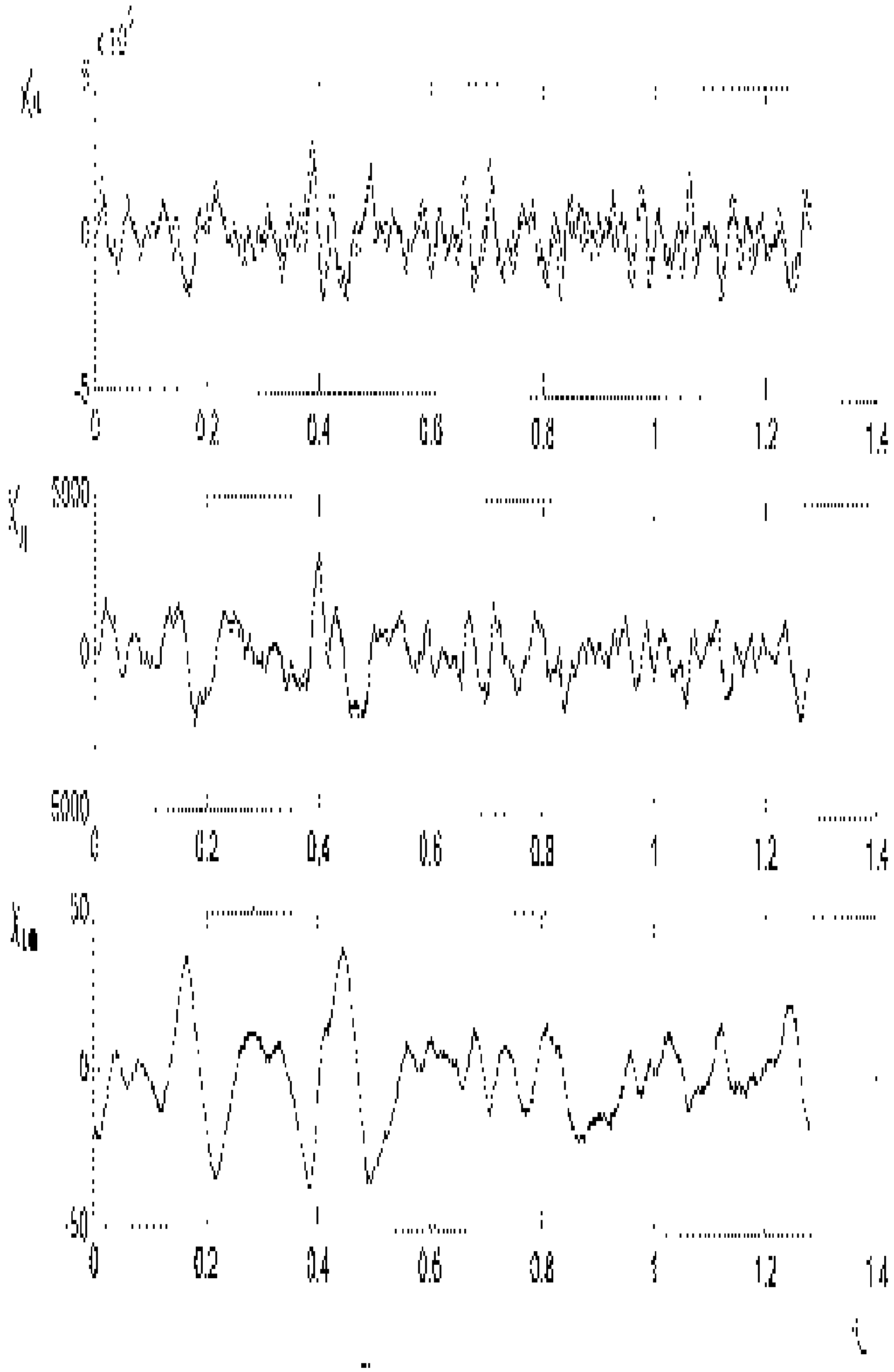


Figure 1c

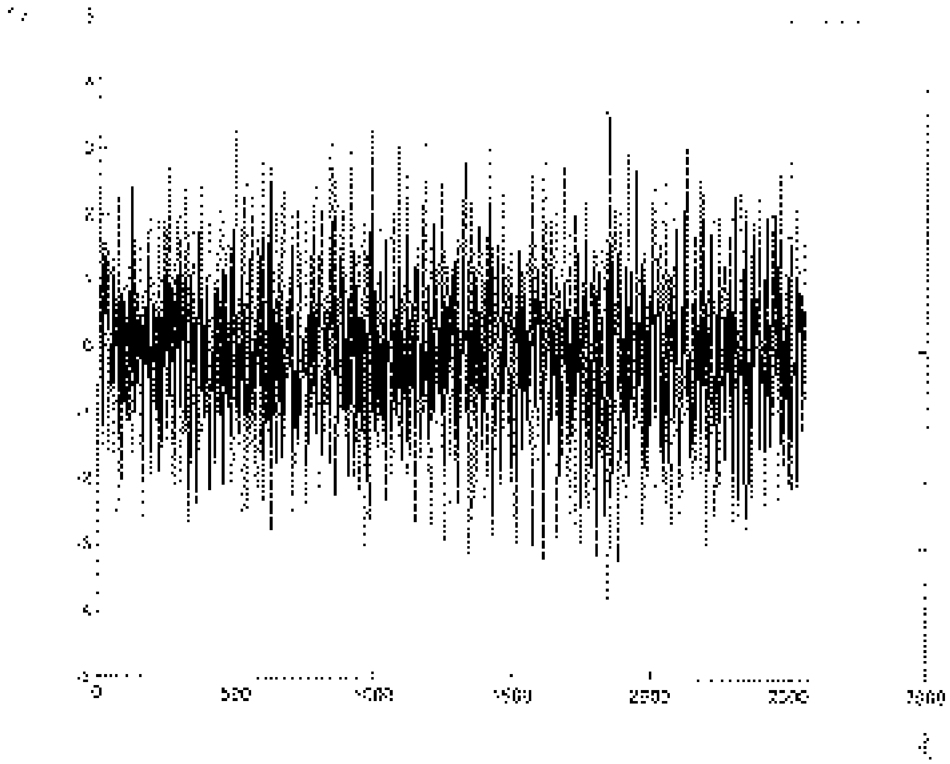


Figure 2

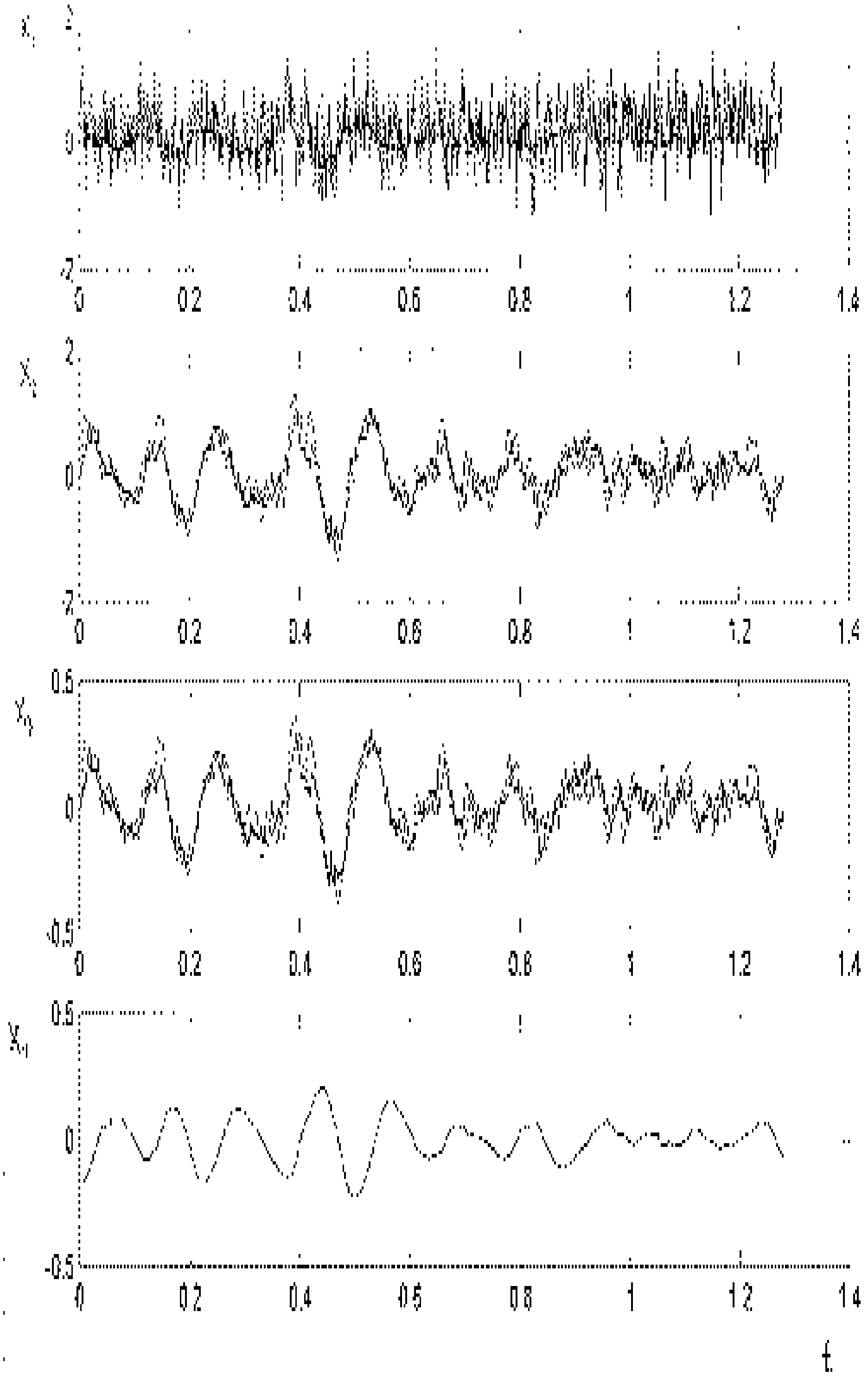


Figure 3a

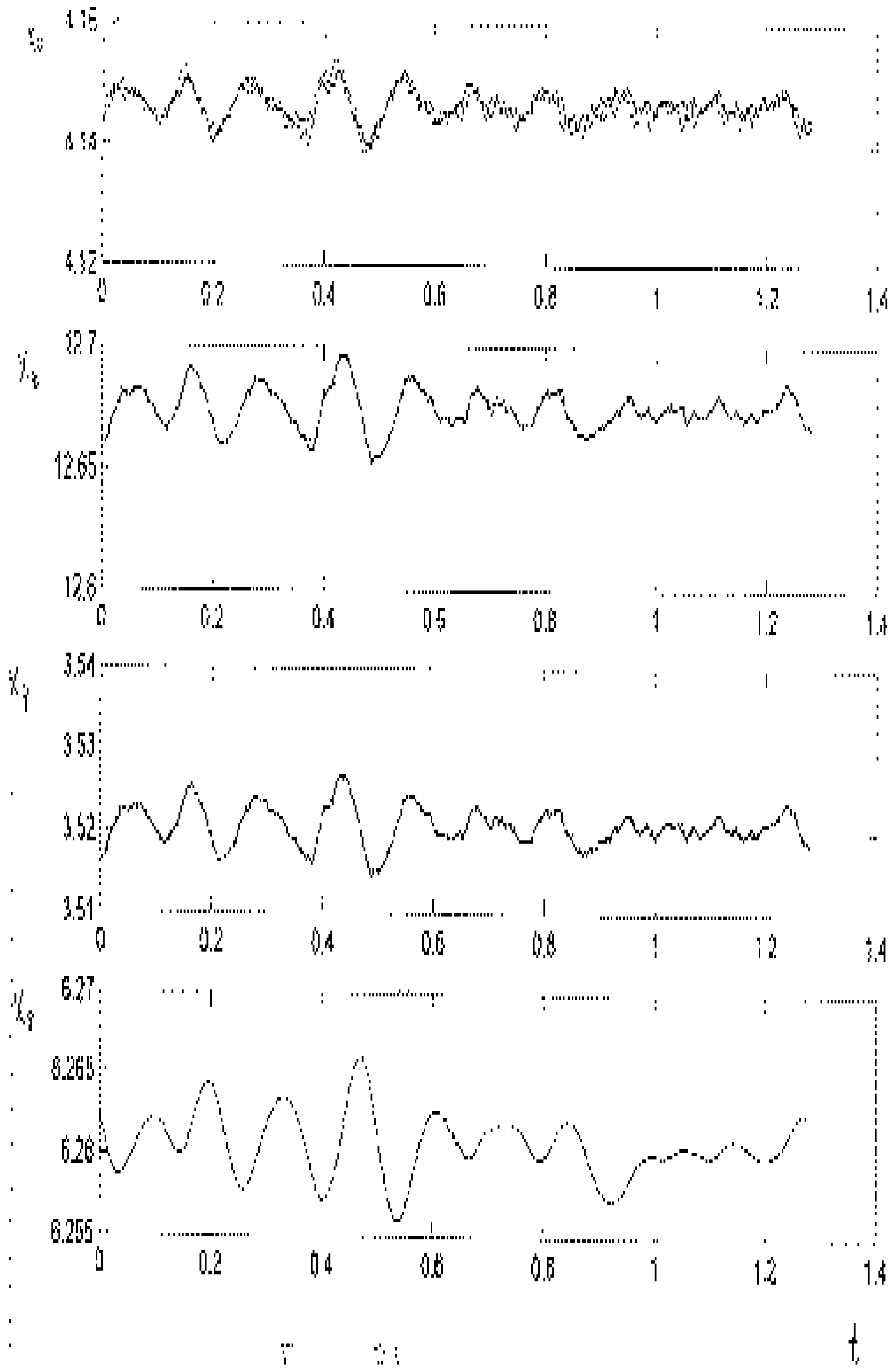


Figure 3b

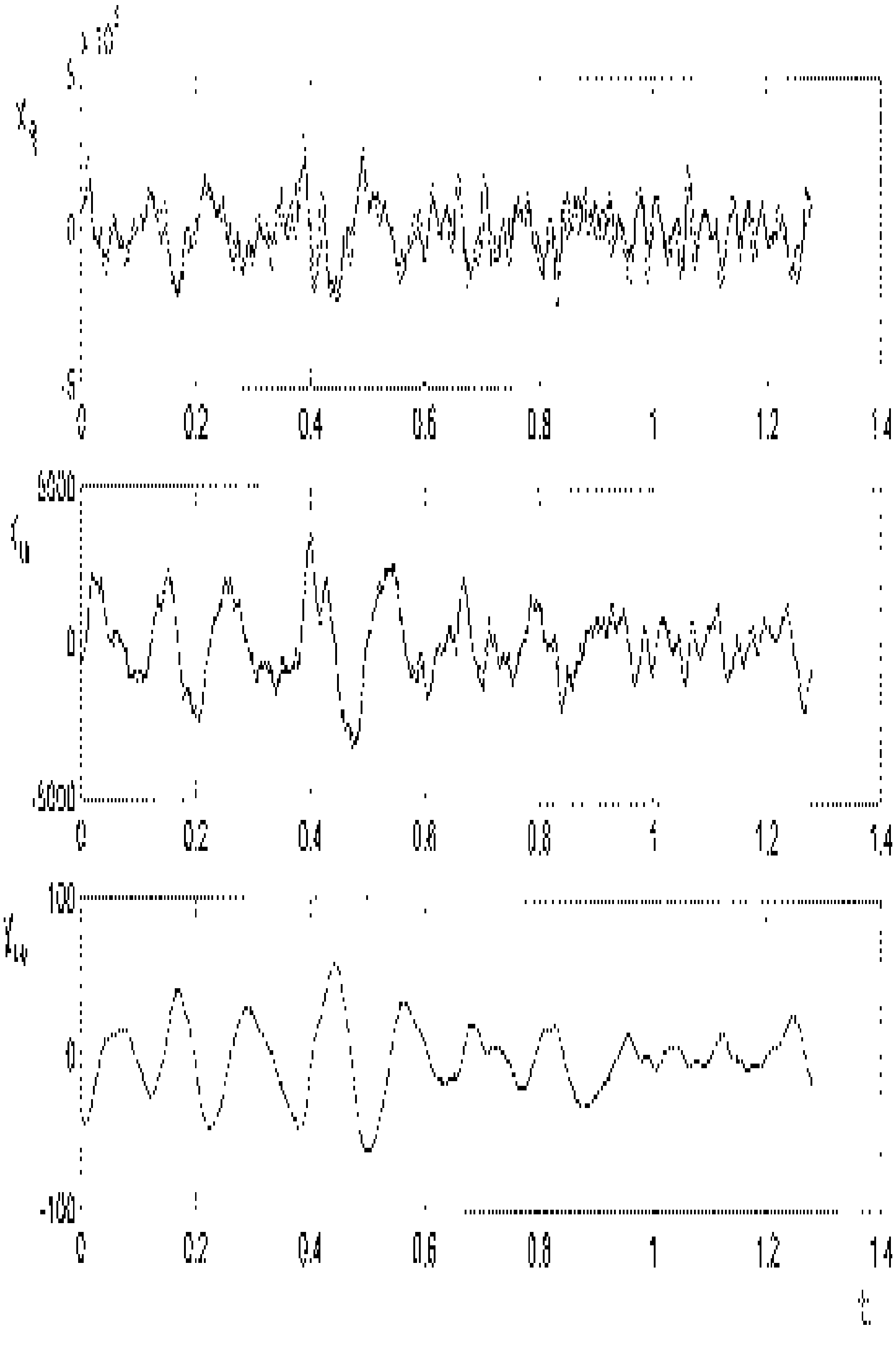


Figure 3c

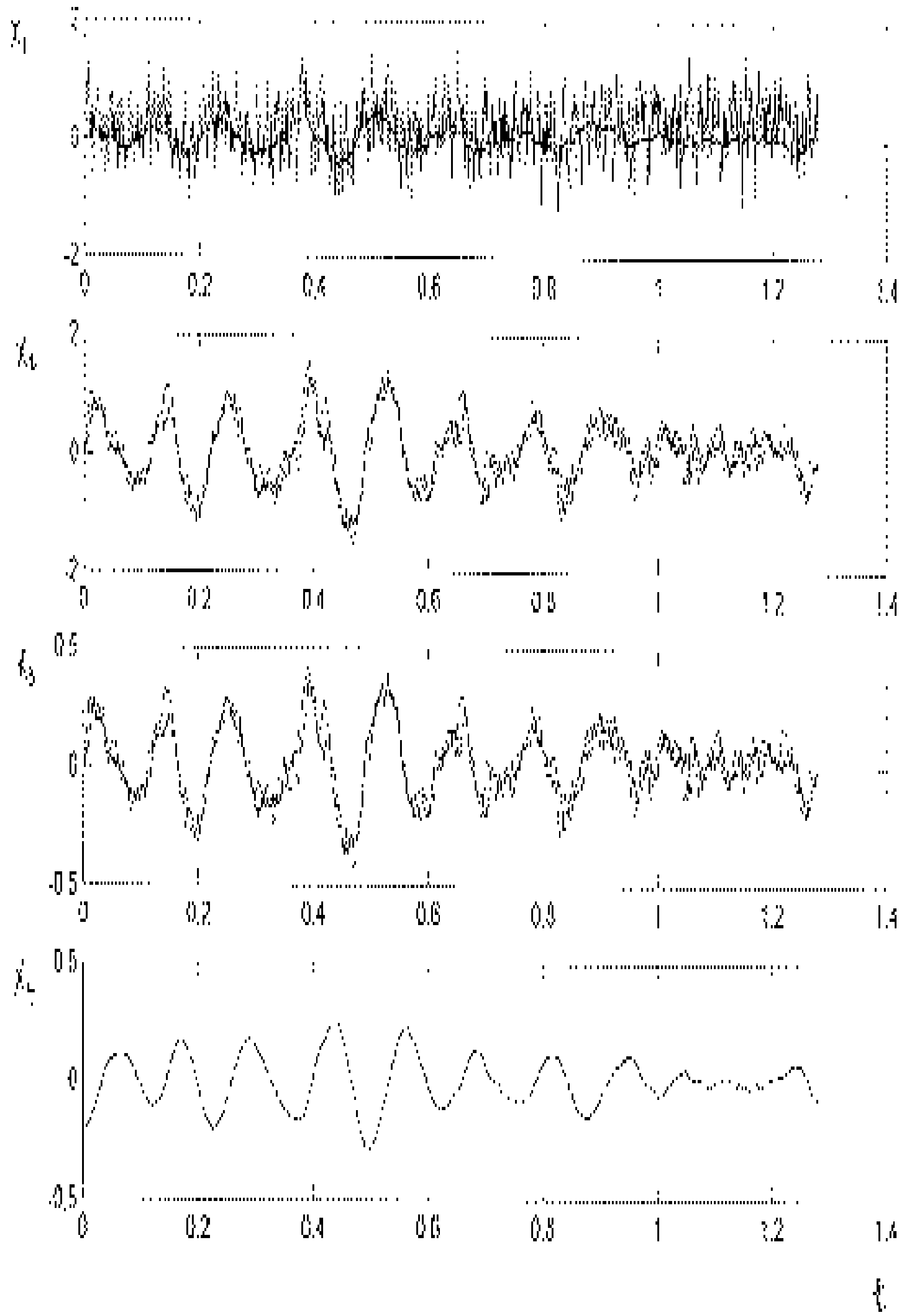


Figure 4a

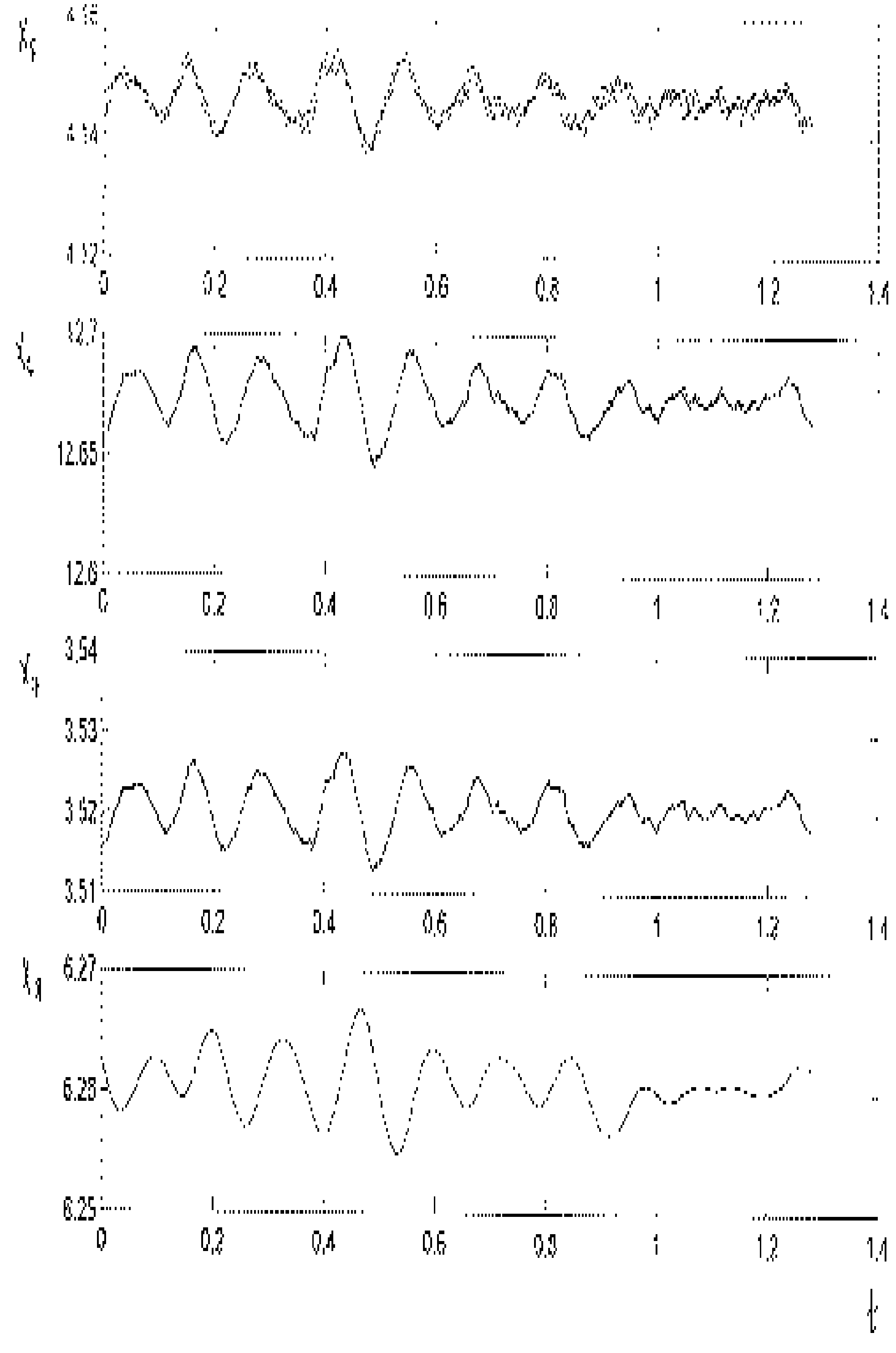
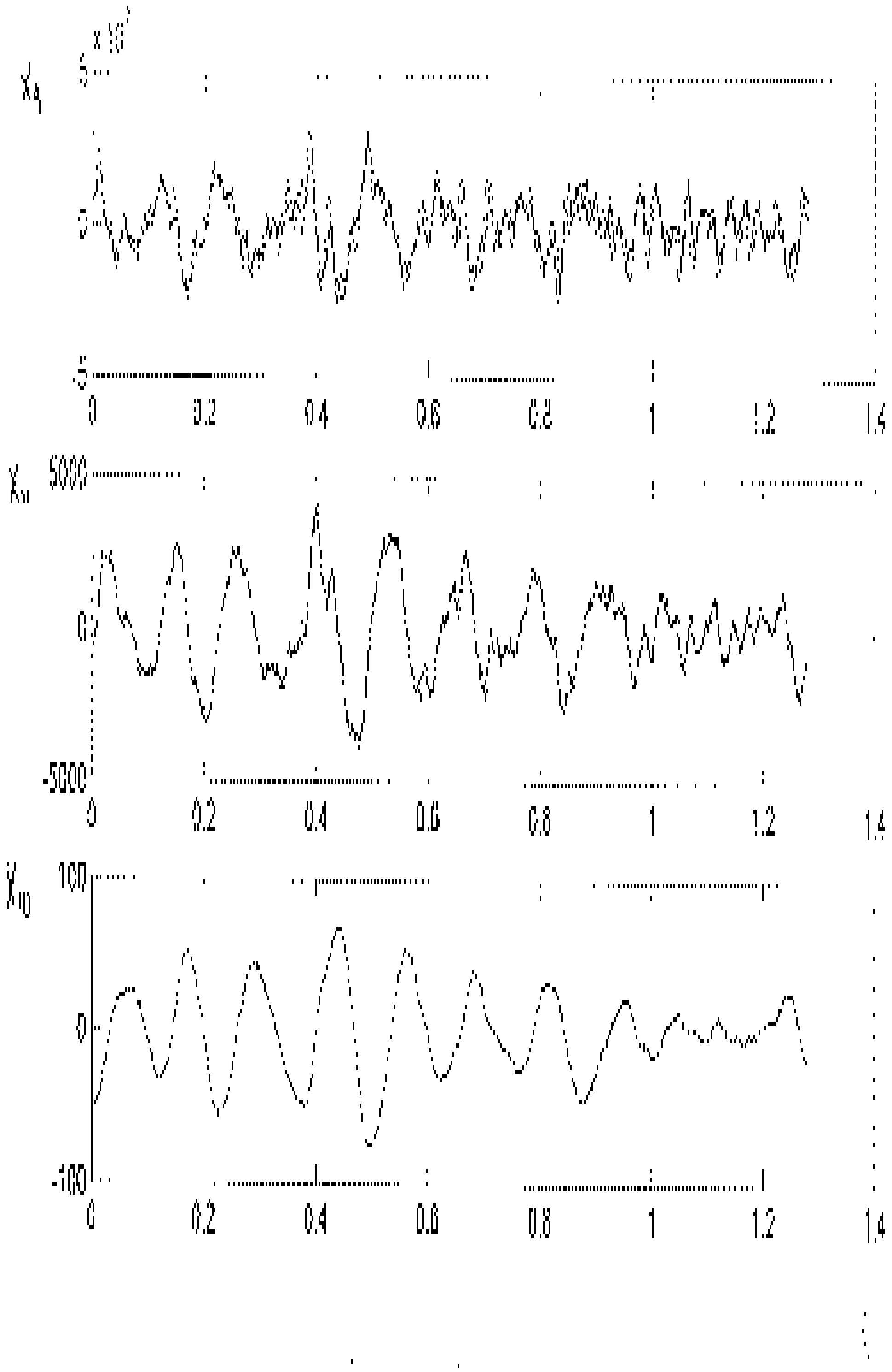


Figure 4b



X_4
x 10
5
0
-5
0
0.2
0.4
0.6
0.8
1
1.2
1.4
X_{11}
5000
0
-5000
0
0.2
0.4
0.6
0.8
1
1.2
1.4
$\dot{X}_{10}$
100
0
-100
0
0.2
0.4
0.6
0.8
1
1.2
1.4

Figure 4c

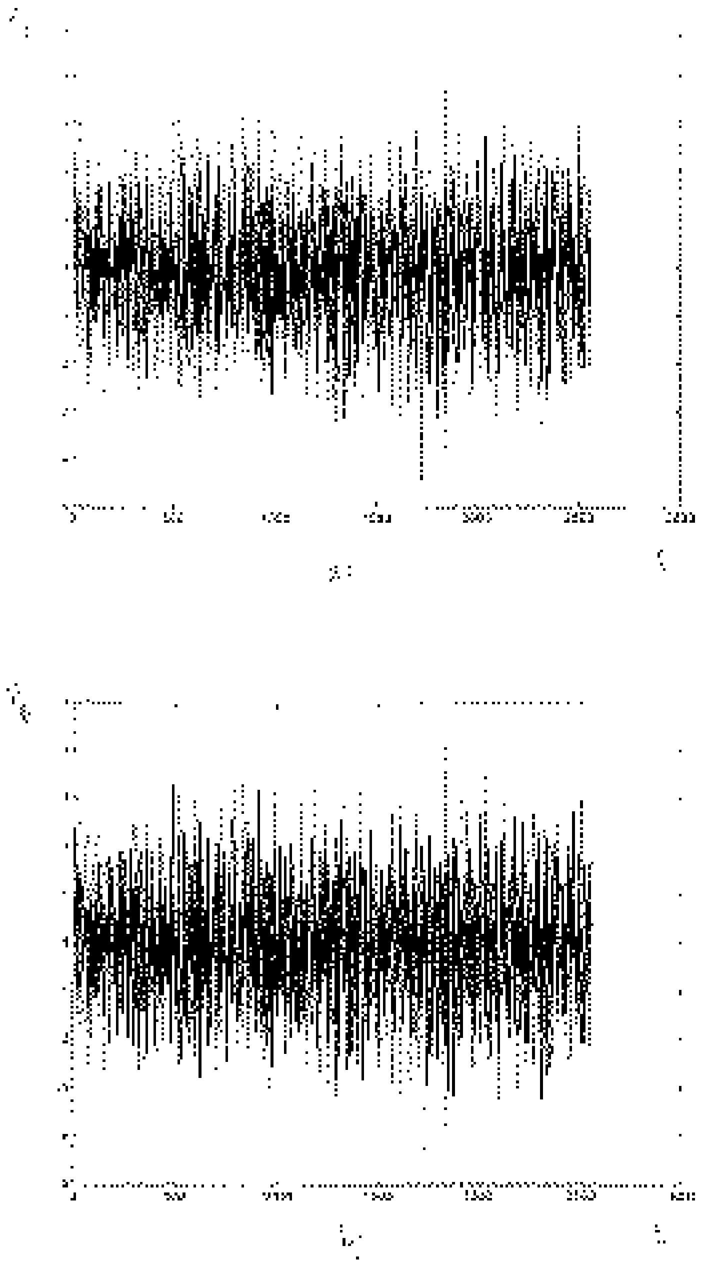


Figure 5

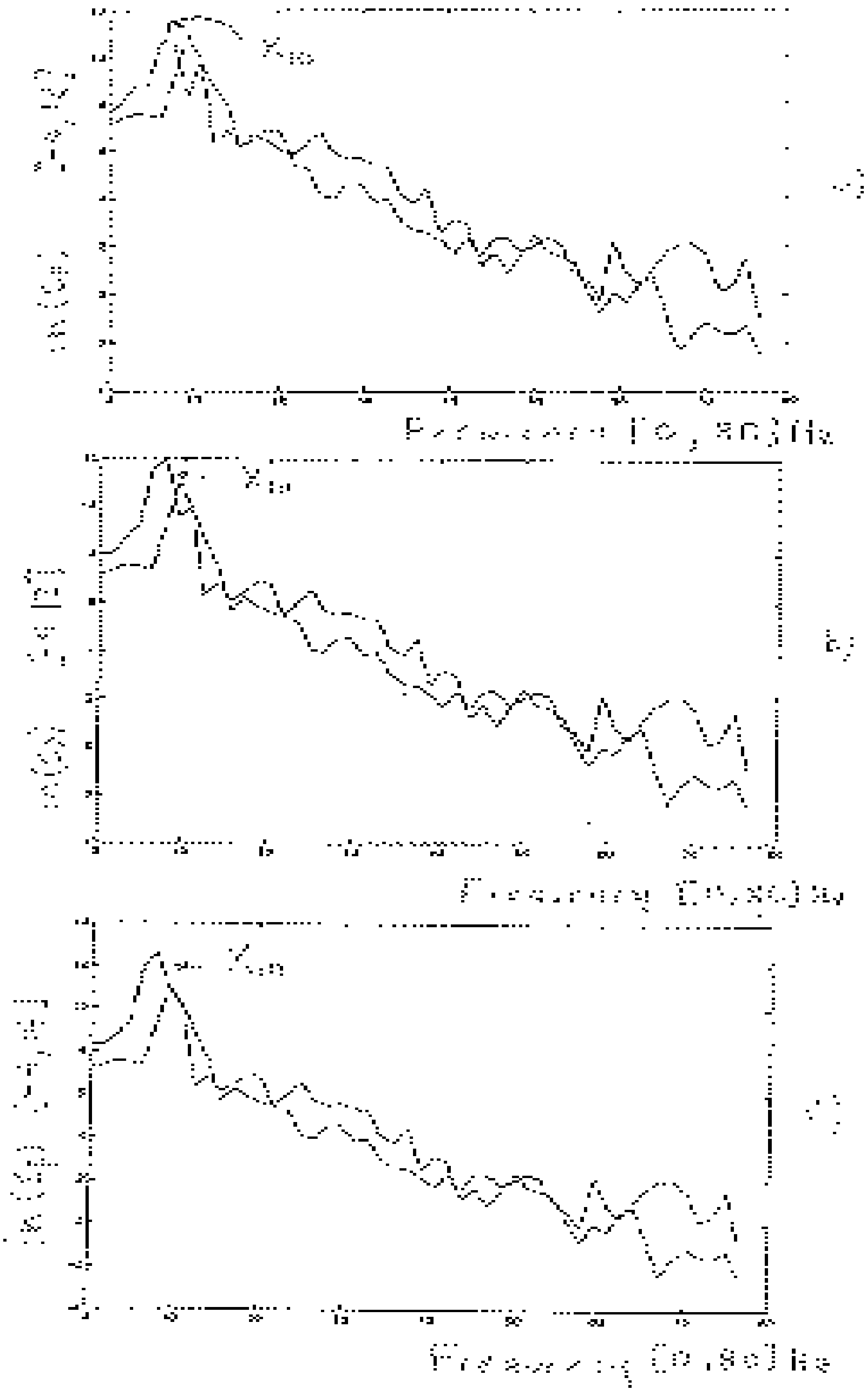


Figure 6