Recursive Estimation of the Continuous-Time Process Parameters

Canudas de Wit, Carlos

1985

Document Version:
Publisher's PDF, also known as Version of record

CODEN: LUTFD2/(TFRT-7290)/1-22/(1985)

Recursive estimation of the continuos-time process parameters

Carlos Canudas de Wit

Department of Automatic Control
Lund Institute of Technology
December 1985

Department of Automatic Control
Lund Institute of Technology
P.O. Box 118
S-221 00 Lund Sweden

Document name
ITERNAL REPORT

Date of issue
December 1985

Document Number
CODEN: LUTFD2/(TFRT-7290)/1-22/(1985)

Author(s)
Carlos Canudas de Wit

Supervisor

Sponsoring organisation

Title and subtitle
Recursive estimation of the continuous-time process parameters

Abstract

The problem of process-parameter identification is considered. Filters are used to create signals related to the derivatives of the input-output process signals. The estimation scheme is implemented by sampling the filtered signals and using a recursive least squares algorithm (RLS). The Choice of filter leads to differents parameter convergence properties. Conditions for parameter convergence are established in terms of the frequency content of the input signal. The relation between choice of filter, sampling time selection and quality of the estimates is discussed and exemplified with simulation examples.

Key words

Classification system and/or index terms (if any)

Supplementary bibliographical information

ISSN and key title

ISBN

Language
English

Number of pages
22

Recipient's notes

Security classification

RECURSIVE ESTIMATION OF THE CONTINUOUS-TIME PROCESS PARAMETERS

*Carlos Canudas de Wit

Department of Automatic Control
Lund Institute of Technology
Box 118, S-221 00 Lund
Sweden

ABSTRACT

The problem of process-parameter identification is considered. Filters are used to create signals related to the derivatives of the input-output process signals. The estimation scheme is implemented by sampling the filtered signals and using a recursive least squares algorithm (RLS). The Choice of filter leads to differents parameter convergence properties. Conditions for parameter convergence are established in terms of the frequency content of the input signal. The relation between choice of filter, sampling time selection and quality of the estimates is discussed and exemplified with simulation examples.

* Laboratoire d'Automatique de Grenoble. B.P. 46, 38402 Saint Martin d'Heres, France. Member of the research group GRECO-Systemes adaptatifs. The Author's work is sponsored by CONACYT-MEXICO.

CONTENTS

Introduction

Identification of the continuous-time process parameters was the initial goal in the earliest work on process identification. The rapid development of digital computers and control theory made the identification of discrete time models and the use of recursive identification techniques more attractive. However, much of the usable a priori process knowlegde is lost after the discretization procedure which is strongly dependent on the sampling conditions, i.e. the mapping between continuous-time and discrete-time process-parameters. Contimuous-time process parameters identification gives the following advantages: 1) More usable a priori knowledge of the process, e.g. order and form of the differential equations, order of magnitude of model-parameter values and known values for some of the parameters. 2) Independent selection of the sampling rate for identification and control purposes.

In order to fit different process descriptions a generalized model was proposed before see, Eykoff (1974), where the signals needed to generate the model error were created by dynamic operators acting on the input-output process signals. In particular when these operators are chosen as low pass filters, it is possible to create signals which are related to the derivatives of the input-output process signals and to estimate the coefficients of the continuous-time differential process representation. For previous work using this technique in identification and self-tuning control see Young (1965, 1969). Stability studies of control schemes using the previous ideas can be found in Johansson (1983), Pernebo (1978) and more recently in studies of 2-D systems by Johansson (1985).

It is clear that the choice of the filters influences the properties of the parameter estimation scheme, e.g, the kind of information produced, the speed of convergence to the solution, the interaction of parameter estimates and the usable a priori process knowledge. In this report we place particular emphasis on the study of the parameter convergence in a deterministic environment and on understanding the influence of the filter choice. Plant, filter and estimation algorithms are related to the well-known convergence condition (persistent excitation). The analysis is carried out in the frequency domain as in previous studies of Sastry (1984).

This work combines the multifilter technique, previously described, with a recursive least squares algorithm. The estimation scheme is described in Section 1. A major effort has been put in to understanding the influence of filter choice on the properties of estimation algorithm. Section 2 reviews the parameter convergence conditions and relates these conditions with the frequency content of the input signal and with the filter spectrum. Section 3 discusses the selection of the sampling time and the filter choice. Some simulations illustrating the main ideas are given in Section 4.

1. Estimation scheme

This section aims at formalizing models to be used in the estimation scheme. As it was mentioned before, the introduction of additional filters permits creation of signals that are related to each other as time derivatives. This method is known as "multifilters technique" and allows identification the parameters of the Laplace representation of the process.

MODEL

Let the physical process be described by a finite dimensional differential equation:

$$D^n y(t) + \ldots + a_o y(t) = b_m D^m u(t) + \ldots + b_o u(t) \qquad (1)$$

where, D^i is the i^{th} differential operator, a_i and b_i are slowly variant or fixed coefficients. Equation (1) describes a proper system , $n \geq m$, having a minimal representation. Taking the Laplace transformation of equation (1), we obtain:

$$s^n y(s) + \ldots + a_o y(s) + Y(0) = b_m s^m u(s) + \ldots + b_o u(s) + U(0) \qquad (2)$$

where Y(0) and U(0) are the initial conditions:

$$Y(0) = \sum_{i=0}^{n-1} \sum_{j=i+1}^{n} y^i(0) s^{n-j} \quad \text{and} \quad U(0) = \sum_{i=0}^{m-1} \sum_{j=i+1}^{m} u^i(0) s^{m-j}$$

$y^i(0)$ and $u^i(0)$ are the initial conditions of the i^{th} derivative of y(t) and u(t), respectively. Let F(s) be a dynamic operator acting on each side of equation (2) and define:

$$y_i(s) \triangleq s^i F(s) y(s) \quad \forall \quad i \in [0,n] \quad \text{and} \quad Y_f(s) \triangleq F(s) Y(0)$$

$$u_j(s) \triangleq s^j F(s) u(s) \quad \forall \quad j \in [0,m] \quad \text{and} \quad U_f(s) \triangleq F(s) U(0)$$

Then equation (2) becomes:

$$y_n(s)+ \ldots+a_o y_o(s)+Y_f(0)=b_m u_m(s)+ \ldots+b_o u(s)+U_f(0) \qquad (3)$$

The time representation of (3) can be obtained by taking the inverse Laplace transform, (L^{-1}), as follows:

$$y_n(t)+ \ldots+a_o y_o(t)+L^{-1}\{Y_f(0)\}=b_m u_m(t)+ \ldots+b_o u_o(t)+L^{-1}\{U_f(0)\} \qquad (4)$$

It can be observed by inspection that the initial conditions in equation (4) will vanish as time increases provided that F(s) be chosen as a stable operator. Equation (4) becomes then an equivalent representation of process description (1). The above structure and the filter signals will be used in the estimation scheme. The creation of the filtered signals is described in the following procedure.

IMPLEMENTATION OF THE FILTERED SIGNALS

The choice of the operator F(s) depends on the type of application and purpose of the estimation scheme. Assuming the interest in models representing the process accurately at low frequencies, it seems intuitive to use a low-pass filter as operator F(s) with the obvious intention of eliminating the high-frequency information. In the sequel, $F(s)=c^l/(s+c)^l$; however, other structures are also possible. Further discussions on the choice of F(s) is refered to Section 3.

The implementation of the filtered signals $y_i(t)$ and $u_i(t)$, can be done in different ways. A straight-forward way is individual implementation of each signal by passing u(t) and y(t) through $s^i/F(s)$. However, it is also possible to construct the signal sets, $y_i(t)$, $u_i(t)$, using only low-pass filters and their linear combinations. This is achieved as follows.

Let us first analyze the case where u(t)=0. Then a system, having y(t) as the input and $y_i(t)$ as the output can be easily constructed:

$$\begin{bmatrix} \dot{x}_1 \\ \vdots \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} -c & c & 0 & & 0 \\ 0 & -c & \ddots & & \\ & \ddots & \ddots & \ddots & 0 \\ & & & -c & c \\ 0 & & & 0 & -c \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ c \end{bmatrix} y \qquad (5)$$

$$\begin{bmatrix} y_o \\ y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}_f = \begin{bmatrix} 1 & & & & \\ -c & c & & & \\ c^2 & -2c^2 & c^2 & & \\ & & & \ddots & \\ (-c)^n & \cdots & \cdots & \cdots & \ddots \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ c^n \end{bmatrix} y \qquad (6)$$

In a compact representation the set (5,6) can be rewritten as:

$$\begin{aligned} \dot{x} &= A_1 x + B_1 y \quad ; \quad A_1 \in \mathbb{R}^{n,n} \quad B_1 \in \mathbb{R}^{n} \\ y_f &= C_1 x + D_1 y \quad ; \quad C_1 \in \mathbb{R}^{n+1,n} \quad D_1 \in \mathbb{R}^{n+1} \end{aligned} \qquad (7)$$

The states x_i are given by $L^{-1}\{x_i(s)\}$, where

$$x_i(s) = \frac{c^{n-i}}{(s+c)^{n-i}}\, y(s) \qquad \forall\ i \in [0,n-1] \qquad (8)$$

and the filtered signals $y_i(t)$ can be obtained by a linear combination of the states x_i, see equation (6). Notice that the coefficients of the triangular matrix C_1 follows the Pascal triangle expansion times $(-c)^i$, i=0, ...,n. Proceeding as before, the general case is solved by implementing the following system with inputs u(t), y(t) and outputs $y_i(t)$, $u_i(t)$.

$$\begin{bmatrix} \dot{x} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} A_1 & \underline{0} \\ \underline{0} & A_2 \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} + \begin{bmatrix} B_1 & \underline{0} \\ \underline{0} & B_2 \end{bmatrix} \begin{bmatrix} y \\ u \end{bmatrix} \qquad (9)$$

$$\begin{bmatrix} y_f \\ u_f \end{bmatrix} = \begin{bmatrix} C_1 & \underline{0} \\ \underline{0} & C_2 \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} + \begin{bmatrix} D_1 & \underline{0} \\ \underline{0} & D_2 \end{bmatrix} \begin{bmatrix} y \\ u \end{bmatrix} \qquad (10)$$

The matrices A_2, B_2, C_2, D_2, have the same form as the matrices defined in (7) but defined in the spaces (n,n), (n,1), (m+1,n) and (m+1,1) respectively. Figure 1 shows the implementation scheme.

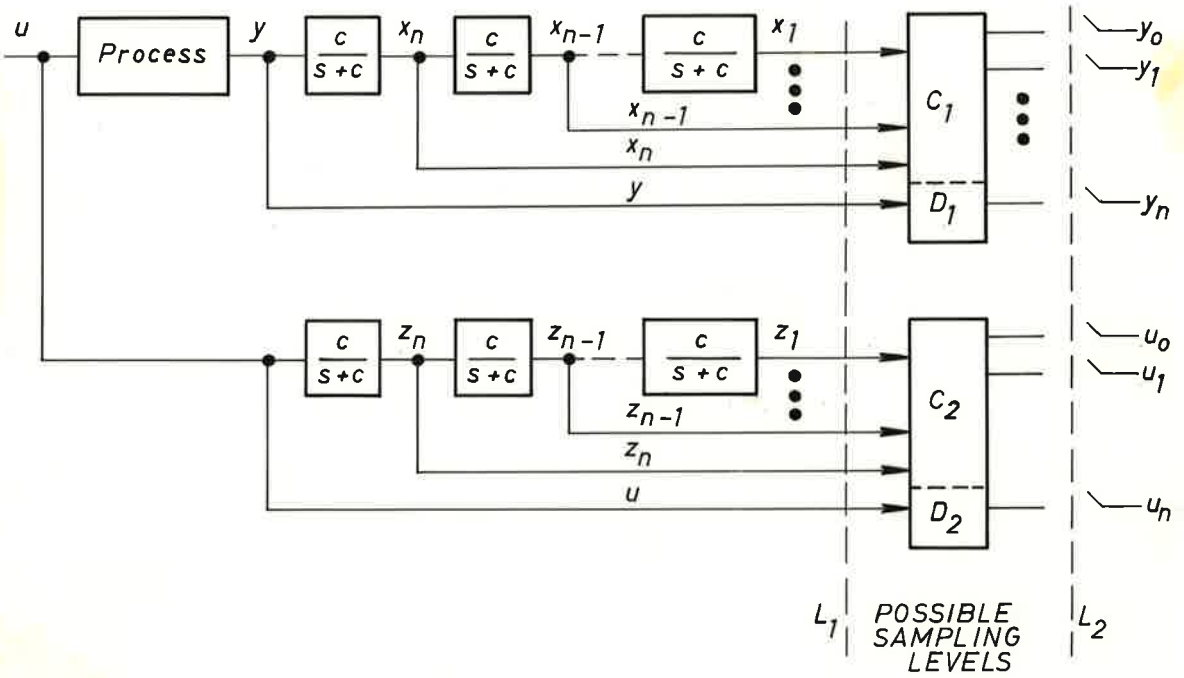

Figure 1. Implementation scheme for multiple-filter technique.

Remark. The signals at level L_1 are easy to create; a simple battery of low pass filters is needed. The signals $y_i(t)$ and $u_i(t)$ can be generated numerically in the computer.

ESTIMATION ALGORITHM.

Differently from previous works, Young (1969, 1969) and Eykoff (1974), the estimation scheme will be based on a recursive algorithm rather than on data recollection and its respective information-matrix inversion. This hybrid scheme allows us to deal with slowly time-variant process and to use the computer potentialities. The parametrization of the model (4) leads to the next compact form:

$$y_n(t) = \theta^T \phi(t) \tag{11}$$

where

$$\theta^T \triangleq [-a_o, \ldots, -a_{n-1}, b_o, \ldots, b_m]$$
$$\phi(t) \triangleq [y_o(t), \ldots, y_{n-1}(t), u_o(t), \ldots, u_m(t)]$$

The observation vector contains continuous-time information that will be sampled at t=0, h, 2h, ..,(named k instants) in order to implement a discrete RLS estimation algorithm. Then the model error will be defined as:

$$e(k) \triangleq y_n(k) - \theta^T \phi(k) \tag{12}$$

The estimation algorithm is obtained by minimizing a model error function (12).

$$J(\hat{\theta}(k)) = \frac{1}{k} \sum_{i=0}^{k} [y_n(i) - \hat{\theta}(k)^T \phi(i)]^2 \tag{13}$$

and described by the following set of recursive equations:

$$\hat{\theta}(k) = \hat{\theta}(k-1) + P(k)\phi(k)[y_n(k) - \hat{\theta}(k-1)^T \phi(k)] \tag{14}$$

$$P(k) = P(k-1)\left[I - \frac{\phi(k)\phi(k)^T P(k-1)}{1+\phi^T(k)P(k-1)\phi(k)}\right] \tag{15}$$

with $P(0) = \alpha I$, for $\alpha >> 0$.

2. Parameter convergence

This section studies the parameter convergence of the previous identification scheme. The noise free case is considered in order to simplify as well as elucidate the analysis. The equations used in the following analysis are very similar to those needed to analyse the noise case where the deterministic signals are substituted by ergodic stochastic process. Since our objective is to choose of the filter F and to see its influence on the parameter convergence, it is useful to translate some of the time-domain properties of the RLS algorithm into the frequency-domain. This permits us to give the parameters convergence properties in terms of the input signal's frequency content and to analyze the influence of the filter in terms of its frequency spectrum.

THE PERSISTENT EXCITATION CONDITION.

The persistent excitation condition is a well-known prerequisite to achieving parameter convergence in identification schemes. For the recursive least squares algorithm (deterministic case), this condition is called "persistent excitation" (PE). Its definition is established as follows:

Definition 1. A vector sequence $\phi(k)$, is persistently exciting if:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \phi(k)\phi(k)^T \geq \alpha I \tag{16}$$

where $\alpha > 0$ and k is an integer multiple of the sampling time, h. In the sequel, the subindex (t) will stand for a continuous time signal, and (k) for a discrete time. Samples are made at t = 0, h, 2h, ... and named k instants.

Definition 2. A bounded vector signal $\phi(t) : \mathbb{R}^+ \to \mathbb{R}^n$ is said to have the limit $R_{\phi(t)}(\tau) \in \mathbb{R}^{n,n}$ iff:

$$R_{\phi(t)}(\tau) \triangleq \lim_{T\to\infty} \frac{1}{T} \int_{r}^{r+T} \phi(t)\phi(t+\tau)^T dt \tag{17}$$

exists for $\{\phi(t), t \in [r,r+T]\}$ with the limit uniform in r. For the sampled vector signal $\phi(k)$, $R_{\phi(k)}(\tau)$ exists only for an integer τ multiple of k, named n.

$$R_{\phi(k)}(n) = \lim_{T\to\infty} \frac{1}{T} \sum_{k=r}^{r+T} \phi(k)\phi(k+n)^T \tag{18}$$

It is useful to relate the PE condition (16) to the equation (17). This allows interpretation of the time-domain condition (16) in terms of the spectral distribution of the input signal u(t). Notice that equation (16) is given in terms of a discrete vector sequence, $\{\phi(k)\}$, and we intend to relate it to the spectral characteristics of a continuous-time signal, u(t). To review the relation of (16) to the Definition 2, we proceed as in previous work of Body and Sastry (1984).

The similarities between PE condition (16) and $R_{\phi(k)}(0)$ in (18) become evident in the following lemmas.

Lemma 1. Suppose that $R_{\phi(k)}(n)$ exists and $\{\phi(k)\}$ is defined for $k \geq 1$, then $\{\phi(k)\}$ is PE iff $R_{\phi(k)}(0) > 0$.

Proof. Assuming the existence of $R_{\phi(k)}(n)$ for r=1, the equation (16) can be rewritten as $R_{\phi(k)}(0) \geq \alpha I$. The rest of the proof is clear from the fact that all positive definite matrices are bounded from zero by their minimun eigenvalue, and reciprocally, matrices having positive bounds are positive definite. Notice also that, $R(\cdot)$ is symmetrical and has real coefficients, hence real eigenvalues.

□

Lemma 2. Provided that $\{\phi(k)\}$ has the limit $R_{\phi(k)}(n)$, defined at the integer set n = 0, h, 2h, ..., and it is positive semidefinite, then from the Helglotz's Theorem, Burill (1973), there exists $S_{\phi(k)}(\nu)$ such that:

$$R_{\phi(k)}(n) = \int_{-1/2h}^{1/2h} e^{i2\pi n\nu} S_{\phi(k)}(\nu)\, d\nu \tag{19}$$

where $S_{\phi(k)}(\nu)$ is called the spectral density function of the sequence $\{\phi(k)\}$. For n=0, $S_{\phi(k)}(\nu)$ is interpreted as the energy density of $\{\phi(k)\}$, or simply, the inverse

of the Fourier transform of $R_{\phi(k)}(0)$.

The next step is to relate the spectrum of u(t) with the condition (16). Let g be the operator vector relating $\phi(t)=g*u(t)$, and $G(i\nu)$ its transfer function.

$$\phi(s) = H(s)F(s)u(s) = G(s)u(s) \tag{20}$$

$$H(s)^T = \left[\frac{N}{D} s^{n-1} \ \ldots, \ \frac{N}{D}, \ s^m \ \ldots, 1 \right] \tag{21}$$

where N/D is the process transfer function and F is the dynamic operator defined in Section 1. Assume that u(t) has a bounded power spectrum, and from the filter theory we have:

$$S_{\phi(t)}(\nu) = H(i\nu)H^*(i\nu)^T |F(i\nu)|^2 S_{u(t)}(\nu) \tag{22}$$

Where $S_{u(t)}(\nu)$ is the spectral density function of u(t). Due to the sampling process, the spectrum of $\phi(t)$ is reproduced at frequencies which are multiple of 1/h.

$$S_{\phi(k)}(\nu) = \sum_{l=-\infty}^{\infty} S_{\phi(t)}(\nu+l/h) \tag{24}$$

Combination of Lemmas 1 and 2 and Equations 22 and 23 allows expression of the time-domain condition (16) in the frequency domain as:

$$R_{\phi(k)}(0) = \int_{-1/2h}^{1/2h} \sum_{l=-\infty}^{\infty} H(i\bar{\nu})H^*(i\bar{\nu})^T |F(i\bar{\nu})|^2 S(\bar{\nu})_{u(t)} d\nu \tag{25}$$

where $\bar{\nu} = \nu+l/h$. Choosing h small enough, to avoid frequency spectrum overlapping, the summation above can be limited to l=0. Then from Lemma 1 and equation (24), $\{\phi(k)\}$ is PE iff,

$$R_{\phi(k)}(0) = \int_{-1/2h}^{1/2h} H(i\nu)H^*(i\nu)^T |F(i\nu)|^2 S(\nu)_{u(t)} d\nu > \alpha I \tag{26}$$

is satisfied. Now, call $u_f(t)$ the signal resulting from $u_f(t)=f*u(t)$, and $S_{uf(t)}(\cdot)$ their spectral density function, the next theorem follows.

THEOREM 1.

The sequence $\{\phi(k)\}$ is PE and hence $\hat{\theta}(t) \rightarrow \theta^*$, iff $S_{uf(t)}(\nu)$ is not concentrated in $K < p$ support points, where p is the parameter number $p = n+m+1$ and

$$S_{uf(t)}(\nu) = |F(i\nu)|^2 S_{u(t)}(\nu) \qquad \forall \ \nu \in [-1/2h \ , \ 1/2h]$$

Proof. Assume first that $S_{uf(t)}(\nu)$ is concentrated in $K < p$ support points, ν_j, within the interval $[-1/2h, 1/2h]$, that is,

$$S_{uf(t)}(\nu) = \sum_{j=1}^{K} \delta(\nu_j) \tag{27}$$

Then, Equation (26) combined with (27) takes the following form:

$$R_{\phi(k)}(0) = \sum_{j=1}^{K} H(i\nu_j)H^*(i\nu_j)^T \tag{28}$$

Since the combination of $K<p$ symmetrical matrices of rank 1 in $\mathbb{R}^{p,p}$ is a singular matrix, $\{\phi(k)\}$ is not PE.

For second part of the proof,see Goodwing and Sin (1984), assume that $S_{uf(t)}(\nu)$ is not zero at more than p points, but there existes a non-zero vector, λ, such that

$$r = \lambda^T R_{\phi(k)}(0)\lambda = \int_{-1/2h}^{1/2h} |\lambda^T H(i\nu)|^2 \, S(\nu)_{uf(t)} d\nu = 0 \tag{29}$$

Hence $H(i\nu)\lambda^T$ vanishes at each support point $\nu = \nu_j$, that is

$$\lambda_1 N(s)s^{n-1} + \ldots + \lambda_n N(s) + \lambda_{n+1} D(s)s^m + \ldots + D(s)\lambda_p = 0 \ , \tag{30}$$

for all $s = i\nu_j$, $j=1, \ldots,p$. Since deg $D(s) = n \geq$ deg $N(s) = m$, there are not more than $m+n = p-1$ roots satisfying the above polynomial equality; this contradicts the existence of λ and proves the theorem.

□

The next corollary reformulates Theorem 1 in terms of the frequency content of u(t).

Corollary 1. $\{\phi(k)\}$ is WPE and hence $\hat{\theta}(t) \rightarrow \theta^*$, iff $S_{u(t)}(\nu)$ is not concentrated in K < p support points and the spectrum of the filter F is non-zero at those points.

From the definition of $S_{uf(t)}(\nu)$, it is clear that if $S_{u(t)}(\nu)$ is concentrated in K support points at ν_j, then $S_{uf(t)}(\nu)$ will also be concentrated at the same frequencies provided that $|F(i\nu_j)|^2$ be non-zero for all j=1, ..,p.

The previous results are in agreement with the expected "richness condition" on the input signal. Notice that, from the above corollary, the asymtotic properties of the RLS algorithm will not be affected by the filter F(s), unless its spectral density be null at the frequencies ν_j's. This is not realizable since it implies the implementation of "ideal filters", i.e. filter with a perferct spectral window. However, the selection of F(s) is critical when more realistic cases are considered, i.e. unmodeled dynamics and presence of noise. Also, the estimation transient will strongly depend on the choice of F(s). Further discussion is defered to Section 4, where some examples are simulated to analyse the transient behavior. The next section discusses the filter selection and its relation to the sampling period and the process band-width.

3. Filter selection

This section discusses the selection of the filter F. As is intuitively expected the choice of the spectrum of F is closely connected with the sampling time and the process band-width. The choice of F is derived from the cost function in equation (13) expressed in the frequency domain as,

$$V(\theta)= \int_{-1/2h}^{1/2h} \left| \frac{N(s)}{D(s)} \right| \left[1 + \sum_{1}^{n+1} a_j s^j \right] - \sum_{0}^{m+1} b_j s^j \Big|^2 |F(s)|^2 S_{u(t)}(\nu)\, d\nu \qquad (30)$$

with $s=i\nu$. See, Wahlberg and Ljung (1984), for further details. It is clear for the above equation and the discussion in Section 1, that the operator F has a double purpose; it creates the equivalent derivatives of the input-output process signals and acts as a weighting function that dictates the relative importance of matching the model to the process at any particular frequency. The estimation algorithm identifies parameters that will enable the model to describe the process accurately in the frequency range where the magnitude of F is high, while representing the process poorly in the frequency range where this magnitude is low.

Then the "ideal" choice of F would be a linear filter with a spectral distribution similar to the desired model spectrum. In practice, however, F is selected as a linear filter of the form $F(s) = c^l/(s+c)^l$, with $l \geq n$. Thus the magnitude of F is constant for frequencies up to c and diminishes thereafter at the rate depending on l. The choice of the above F is an attractive candidate for identifing models at low frequencies.

SAMPLE TIME SELECTION.

Refering to Figure 1, the use of the sampling level L1 imply to sample the states $x_i(t)$, $z_i(t)$ and the input-output signals u(t) and y(t). It is also possible, by chosing $l > n$, to construct the signals set $\{y_i(t)\}$ and $\{u_i(t)\}$ only with the states x_i and z_i for $i = 1, ..,n$. Then according to the sampling theorem the sampling frequency, ν_s, should be chosen as, at least, twice the largest frequency band-width of the

signal set $\{ x_i, z_i\}$, named $\nu_0 = \max B\{x_i, z_i; \forall\ i = 1, ..,n\}$. Its is clear that ν_0 will depend basically on the process band-width and the filter cut-off frequency. Low-frequency models for which we have a rough notion of the process frequency band-width, ν_p, it seems reasonable to choose F as a low-pass filter having a frequency band-width, ν_f, grater than or equal to ν_p, and select the sampling frequency as,

$$\nu_s \geq 2\nu_f \geq 2\nu_p \tag{31}$$

Notice that the lower bound of ν_f in (31) assures the identifiability of the model at frequencies lower that ν_p. For practical reasons it is advisable to select $\nu_f \approx \nu_p$ (reasonable suppression of the undesirable higher frequencies and process parasites).

4. Examples

The following examples aim is to show the influence of the filter F on the performance of the identification scheme. The examples are realized a in hybrid environment (using SIMNON package). The process to be identified is a first order system, described by the following differential equation:

$$\frac{d}{dt}\, y(t) + ay(t) = bu(t) \tag{32}$$

The parameters a and b will be estimated using the technique described in Section 1. F is chosen to be a low-pass filter of the form $F(s) = c^2/(s+c)^2$. The input signal is periodic, $\cos(2\pi\nu_u t)$, with its density energy concentrated on two points; $\pm\nu_u$. A schematic diagram of the filter implementation is shown in Figure 2.

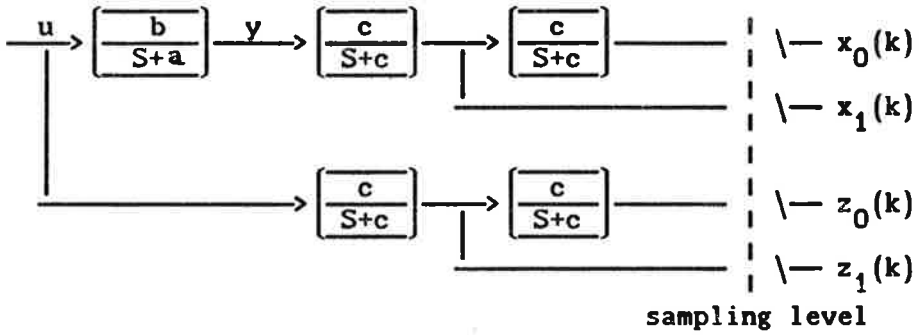

Figure 2. Multifilter implementation for a first order process.

EXAMPLE 1.

This example illustrates the transient behavior of the estimates for different values of c. Sampling frequency is selected large enough to avoid alias phenomena, ν_s=1000 Hz. The process parameters are; a = 10 and b = 12. ν_u is 10 Hz.

The simulation results are shown in Figure 3. It is clear that the estimates converge fast to their optimal values as the value of c is increased. However, the selection of largest values of c are limitated by practical considerations. Large values of ν_f (or similarly large c) imply to increase the sampling frequency ν_s, see equation (31), it will also intensify the influence of undesired high frequency information (presence of noise and unmodeled dynamics). This last case is considered in the next example.

EXAMPLE 2

This example illustrates the influence of F when more realistic situations are considered. The plant is described by the following transfer function;

$$G(s) = \frac{1200}{(s+10)(s+100)} \qquad (33)$$

which is the same as before but with a parasite pole located at 100 Hz. White noise (0, 15%) is also added to the process output. The equation (32) describes the desired model to be matched to the plant (33). The values of ν_s and ν_u are the same as in the previous example.

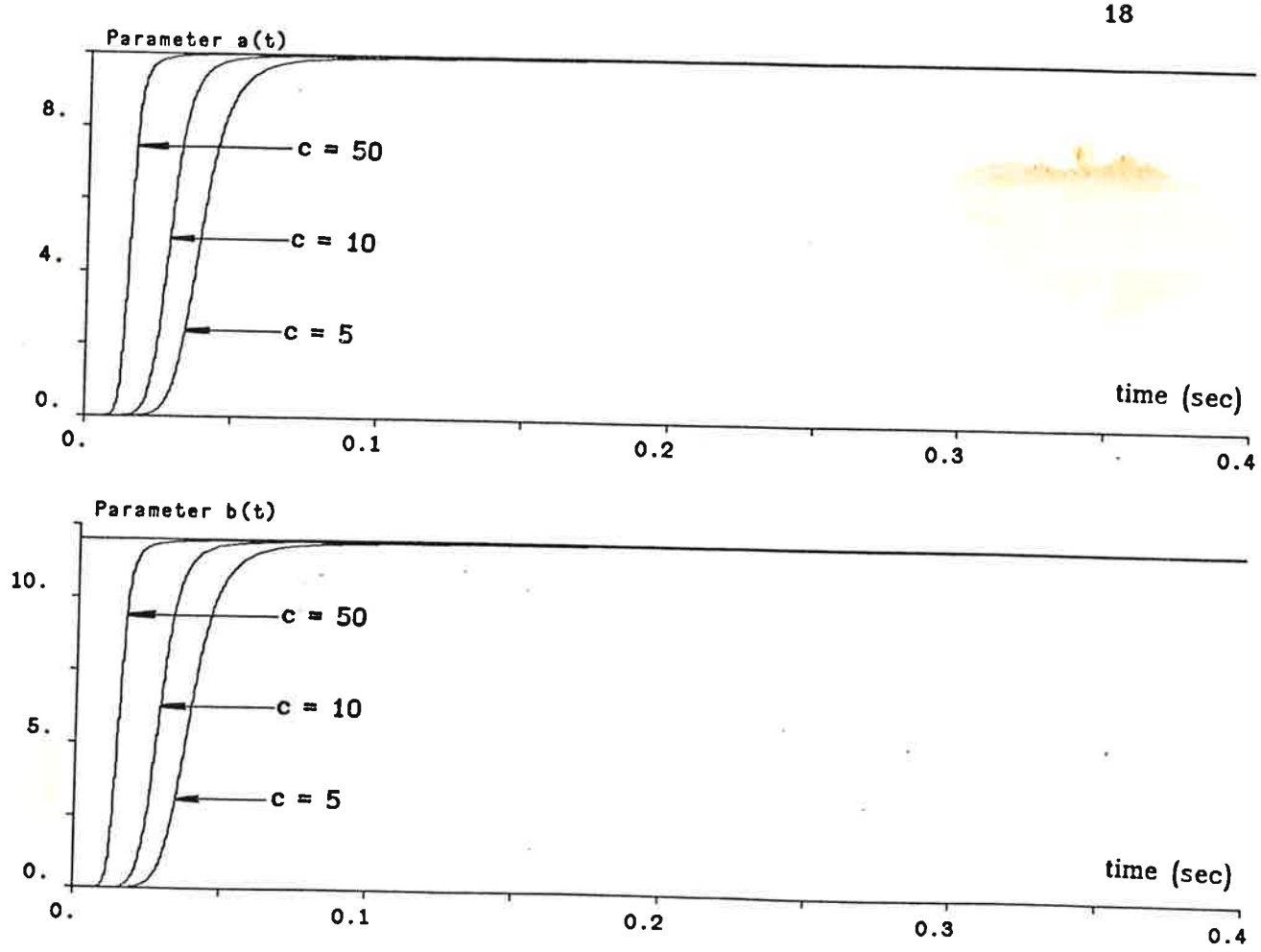

Figure 3. Transient behavior of the estimates $\hat{a}$ and $\hat{b}$ for differents values of the filter cut-off frequency c.

The presence of noise and unmodeled dynamics will produce a bias in the estimation of a and b. The bias levels will certainly depend on the choice of F. Notice that, selecting F as a low-pass filter (c = 10 Hz.), the identified parameters will enable the model to describe the process accurately only in the frequency range 0-10 Hz.

Figure 4 shows that better results are obtained by choosing values of c close to 10 Hz. The identified model has a static gain of 1.22 and a time constant of 9.42 (Process static gain is 1.2 and the dominant time constant is 10). As the value of c is increased the transient behavior of the estimates deteriorates and the bias levels are increased. For c=100, the identified static gain is 1.1 and the time constant is 8.11.

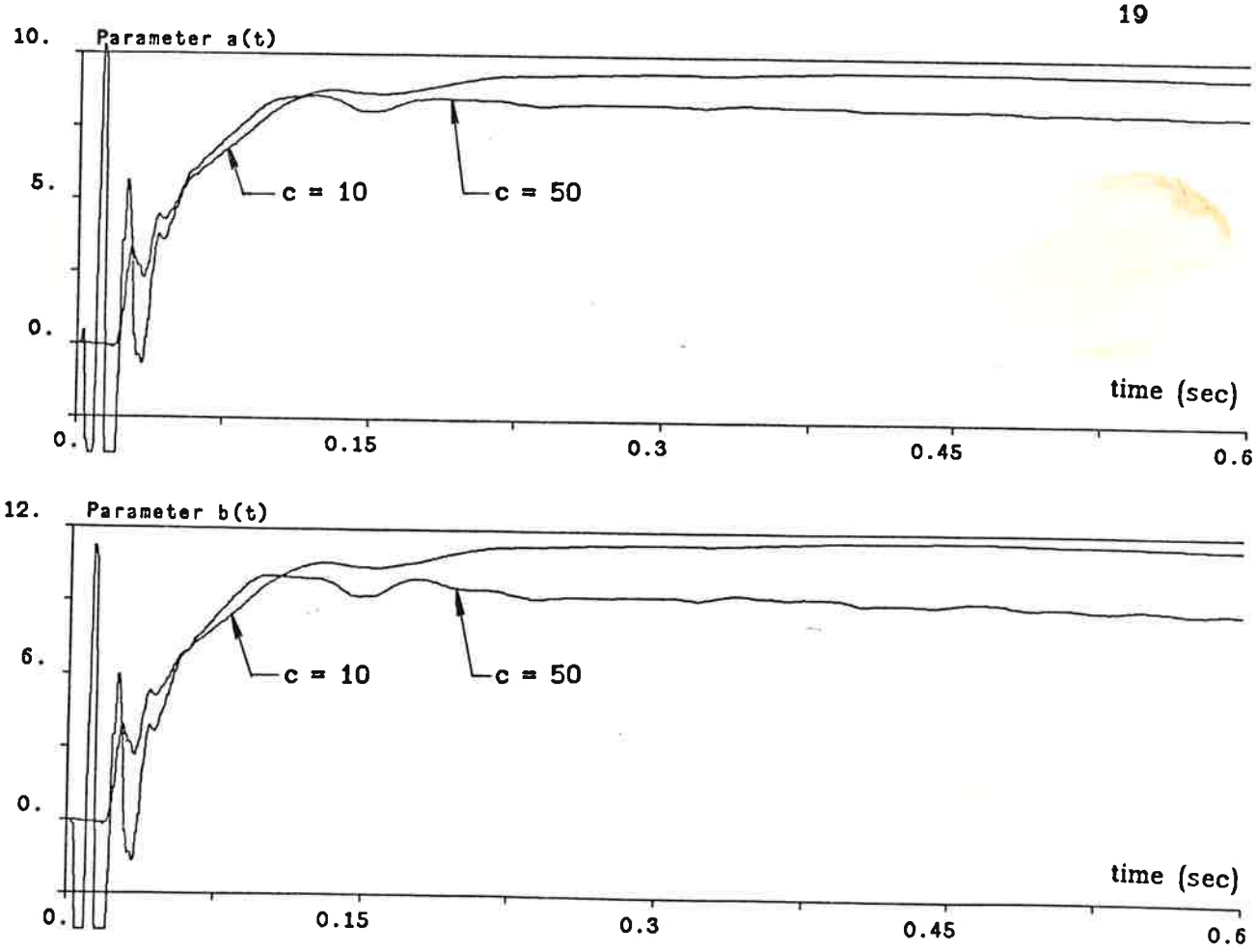

Figure 4. Selecting F in presence of noise and neglected dynamics.

Conclusions

Process parameter identification gives a powerful tool in modeling and control design. Its major advantage is that the set of identified model parameters is directly related to the differential equation that describes the physics of the process. Section 1 has reviewed this technique and described the manner of implementation. The implementation should be realized as simply as possible in order to avoid too many analog elements. The simplest way is using a battery of low-pass filters and sampling at level L1, see Fig. 1.

The conditions on parameter convergence were established in Section 2 in terms of the frequency content of the input signal. These conditions can be sumarized as follows: for input signals having a spectral density $S_{u(t)}(\nu)$, it is sufficient that $S_{u(t)}(\nu)$ has p or more support points in order to archive parameter convergence.

Sample-time selections and alternatives of choice of F were analyzed in Section 3. Section 4 was dedicated to illustrate the effects of the choice of the filter F in two situations. For the deterministic case with adequate process and model orders, the choice of F is not critical and affects only the rate of convergence. In more realistic situations (presence of noise and unmodeled dynamics), the selection of F is more relevant and may change drastically the estimates quality. The ideal choice of F would be a linear filter with a spectral distribution similar to the desired model spectrum.

Other interesting topics remain to be analyzed; i.e. convergence rate as a function of the filter and input spectrum, estimation bias produced by model mismatch, and noise contamination.

Acknoledgements

I would like to thank, Profesor K.J. Åström and all the members of the department of Automatic Control, Lund Institute of Thechnology, for their hospitality and helpful sugestions. Profesor I.D. Landau my adviser at LAG-ENSIEG Fance. The next organizations: CONACYT (Mexico), CEFI and GRECO-Systemes adaptatifs (France), which support the author's work.

References

Boyd, S. and Sastry, S.S. (1984). "Necessary and Sufficient conditions for Parameter Convergence in Adaptive control",IFAC'84, L.V.N. USA.

Eykhoff, P. (1974). "System Identification", Wiley- Interscience publication.

Goodwin, G.C. and Sin, K.S. (1984). Adaptive Filtering Prediction and Control. Prentice-Hall, INC.

Johansson, R. (1983), "Multivariable Adaptive Control", Doctoral Dissertation , CODEN: LUTFD2/(TFRT-1024)/1-207/(1983). LTH, Lund, Sweden.

Johansson, R. (1985), "Estimation and Direct Adaptive Control of Delay-Differential Systems." IFAC'85, York, UK.

Pernebo, L. (1978), "Algebraic Control Theory for linear Multivariable Systems ", Doctoral Dissertation , CODEN: LUTFD2/(TFRT-1016)/1-307/(1978). LTH, Lund, Sweden.

Young, P.C. (1965). "Process parameter estimation and Self-Adaptive Control System (Ed. Hamnond), Plenum Press, New York, pp 118-139.

Young, P.C. (1969). "Applying Parameter Estimation to Dynamics Systems", Control Engng. 16, Oct. 119-125; Nov. 118-124.

Wahlberg, B. and Ljung, L. (1984). "Design variables for bias distribution in transfer function estimation". Proceedings of 23rd CDC, Las Vegas, NV.