IDENTIFICATION FOR

DISTRIBUTED PARAMETER SYSTEMS*

Carlos Silva Kubrusly

ABSTRACT

This thesis considers the parameter identification problem for systems governed by partial differential equations. The various identification methods are grouped into three disjoint classes namely: "Direct Methods", "Reduction to a Lumped Parameter System", and "Reduction to an Algebraic Equation".

The major subject investigated here is concerned with the applicability of stochastic approximation algorithms for identifying distributed parameter systems (DPS) operating in a stochastic environment, where no restriction on probability distributions is imposed. These algorithms are used as a straightforward identification procedure, converge to the real value of the parameters with probability one, and are suitable for on-line applications. In this way, a new identification

---

* This research was supported by CNPq. (Brazilian National Research Council) under Grant No. 3712 / 73.

IDENTIFICATION FOR DISTRIBUTED PARAMETER SYSTEMS*

Carlos Silva Kubrusly

ABSTRACT

This thesis considers the parameter identification problem for systems governed by partial differential equations. The various identification methods are grouped into three disjoint classes namely: "Direct Methods", "Reduction to a Lumped Parameter System", and "Reduction to an Algebraic Equation".

The major subject investigated here is concerned with the applicability of stochastic approximation algorithms for identifying distributed parameter systems (DPS) operating in a stochastic environment, where no restriction on probability distributions is imposed. These algorithms are used as a straightforward identification procedure, converge to the real value of the parameters with probability one, and are suitable for on-line applications. In this way, a new identification

---

* This research was supported by CNPq. (Brazilian National Research Council) under Grant No. 3712 / 73.

method is developed for DPS described by linear models, driven by random inputs, and observed through noisy measurements. The very real case of noisy observations taken at a limited number of discrete points located in the spatial domain is considered. The proposed identification method assumes that a previous system classification has been performed, such that the model to be identified is known up to a set of space-varying parameters, where extraneous terms may be included.

ACKNOWLEDGEMENTS

I wish to thank my supervisor, Dr. Ruth F. Curtain, for her constant encouragement in all stages of this work. Her advice and observations were particularly valuable to me.

I am also grateful to Mr. Luca Braito for his assistance in computer programming.

Special thanks are due to Berenice for her efficient and careful typing of a difficult manuscript.

CONTENTS

INTRODUCTION

The central theme in this thesis is the application of stochastic approximation theory, as a straightforward identification technique, for determining parameters in systems governed by partial differential equations (PDE).

In chapter 1 we begin by presenting some introductory aspects of the general problem of system identification. Fundamental concepts (such as: system characterization, classification and identification, as well as lumped (LPS) and distributed (DPS) parameter systems) are explained in order to make precise what kind of problem will be considered here.

A survey on the DPS identification field is presented in chapter 2. Before reviewing the various approaches used to face the problem, we introduce a new classification for the DPS identification methods. Briefly, these methods can be grouped into three disjoint classes: The first one uses optimization techniques directly on the model that describes DPS. The second class of methods is characterized by reducing the DPS to an equivalent LPS. In a similar way, the methods in the third class reduce the DPS to a set of algebraic equations.

Chapter 3 is concerned with the mathematical concepts and techniques that will be used later for identification purposes. It contains three independent parts: In part I we consider some classes of models for DPS described by PDE. Higher order finite-differences are introduced in part II, where the basic lemmas for model approximation are derived. Relevant aspects of the stochastic approximation theory, as well as some applicable stochastic approximation algorithms in Hilbert space, are presented in part III.

The main results appear in chapter 4. There we propose a new method for identifying space-varying parameters in distributed systems. The DPS is supposed to be operating in a stochastic environment, and no restriction concerning probability distributions is imposed. A class of linear models (where extraneous terms may be included) driven by random inputs and observed through noisy measurements is considered. These measurements are taken at a limited number of discrete points located in the spatial domain. The theory is developed by assuming a one-dimensional spatial domain, but direct extensions to multi-dimensional spatial domains can be obtained as shown in section 4.7. Higher order finite-difference techniques are used to reduce the DPS to an equivalent discrete-time LPS. The parameters are then placed in an explicit form which is suitable for applying recursive identification schemes. In this way, stochastic approximation algorithms (as proposed in chapter 3, part III) are used as a straightforward on-line identification procedure, rather than a simple searching scheme for finding estimates previously obtained by means of any other optimization technique. These algorithms converge to the real value of the parameters with probability one.

Finally, the performance of the identification method is analysed in chapter 5. After a brief summary concerning second-order models, we present three examples dealing with parabolic and hyperbolic PDE. Conclusions and suggestions for further research are also included.

References are listed at the end of the thesis, and grouped according to chapter.

CHAPTER 1

SOME BASIC ASPECTS IN SYSTEM IDENTIFICATION

The general idea of System Identification is a very wide one and different authors dealing with this subject use the term "system identification" in some slightly different ways.

It is not our intention in this study to pose a formal or rigorous definition of system identification. Instead of this, we intend to present an informal and brief introduction on this topic as a starting point for the subsequent chapters.

The three stages of the identification procedure namely, system characterization, system classification and system identification are discussed in section 1.1. The meaning of lumped parameter system and distributed parameter system is explained in section 1.2. Finally, in section 1.3, the problems of system identification and state estimation are discussed and the difference between these two concepts is emphasized.

1.1 - SYSTEM CHARACTERIZATION, CLASSIFICATION AND IDENTIFICATION

One of the first attempts to explain the main concepts involved in System Theory was made by Zadeh [1]. Under the subtitle "Principal Problems of System Theory", he formulated twelve of the most important problems (both from theoretical and practical viewpoints), which are summarized below:

1) System Characterization

2) System Classification

3) System Identification

4) Signal Representation

5) Signal Classification

6) Systems Analysis

7) Systems Synthesis

8) System Control and Programming

9) System Optimization

10) Learning and Adaption

11) Reliability

12) Stability and Controlability

It would be helpful, for the purpose of our study, to add two important problems to the above list:

13) Observability

14) State Estimation (Filtering, Smoothing and Prediction)

As we are interested only in the meaning of the first three of those problems and mainly in the third one, we will discuss these in the light of Zadeh's paper.

<u>System Characterization</u>: "Representation of input-output relationships in mathematical form; transition from one mode of representation to another".

This problem is concerned both with the different ways in which the input-output relationship of a specific system can be represented (i.e., in terms of differential equations, state equations, characteristic functions, frequency response functions, integral operators, etc.), and the forms which these representations assume for several types of systems (i.e., continuous-time, discrete-time, stochastic, deterministic, memoryless, finite-memory, causal, etc.).

System Classification: "Determination on basis of observations of input and output, of one among a specified class of systems to which the system under test belongs".

In the following we will call the "class of systems" by the class of models or simply the class, the elements of a class are obviously called by models or mathematical models, and the "system under test" will be called the system (some authors call it process or plant).

This kind of problem may be stated as follows:

Assume that

i) I is an index set,

ii) C_α, $\alpha \in I$, are classes of models M,

iii) $F = \{C_\alpha;\ \alpha \in I\}$ is a family of these classes.

Suppose we are given a system S and a family F, such that S is characterized by F and belongs[1] to one of its classes, say C_α. The problem is to determine C_α by observing the responses of S to some different inputs.

---

1 The expression "S belongs to C_α" is not a very accurate one. What really belongs to C_α is a specific model, say M^*, which is in some sense "equivalent" to S. The meaning of "equivalent" will be precised later in this chapter.

Example 1. An important particular problem in classification is the following:

Let

i) $I = Z_+ = \{1, 2, 3, \ldots\}$: the set of all positive integers.

ii) C_n be the class of all models M described by a single ordinary differential equation of order n. [2]

iii) $F = \{C_n ; n \in Z_+\}$ be the family of all these classes. That is, the set of all ordinary differential equations of any finite order.

Suppose S characterized by F. Which means: suppose it is known that the system is represented in terms of an ordinary differential equation (this assumption concerns the System Characterization).

The question is: What is its order? In other words, which class $C_n \in F$ does S "belong"? (or better: which class $C_n \in F$ does M*, the "equivalent" model, belong ?)

The problem of finding the class C_α is, sometimes, described as the black box approach. Roughly speaking, this means the determination of the structure (or "topology") of the system, considering it as a perfect "black box". On the other hand, assuming that some

---

[2] All ordinary differential equations belonging to a given C_n (i.e., the models $M \in C_n$) are completely known up to a set of m $(m > n)$ parameters, which are the coefficients of the differential equation. The problem of finding these parameters concerns to the System Identification, and it will be commented later in this section.

class of models, say C_α , is available[3] the determination of one element M in C_α is, sometimes, called the opaque box approach. This is the subject of System Identification which is described below.

System Identification: "Determination, on basis of observations of input and output, of a system within a specified class of system to which the system under test is equivalent".

Observing the nomenclature introduced before (i.e., the meaning of class, model and system) the identification problem may be formulated as follows:

Given a class C_α (with each member of C_α completely characterized), the problem is to determine a model M in C_α which is equivalent to the system S. Briefly: find $M \in C_\alpha$ such that M is equivalent to S.

But what does the term "equivalent" mean in this particular case ?

Assume that

i) W is some space of inputs and w a element of W.

ii) $y_S = y_S(w)$ is the system output and $y_M = y_M(w)$ is the model output, for some pre-selected input w in W.

---

3 This assumption can be based on some "a priori" knowledge of the system's structure. This "a priori" knowledge can be thought as the result of a previous classification. We will make more comments about what we mean by "a priori" knowledge, later in this chapter.

The equivalence is often defined in terms of a cost function J which is a functional of y_S and y_M. That is:

$$J = J(y_S, y_M)$$

The model M* which is <u>equivalent</u> to the system S will be one such that the cost function J is minimized . Symbolically we have

$$M^* \backsim S \iff J(y_S, y_{M^*}) = \underset{M \in C_\alpha}{\text{Min.}}\ J(y_S, y_M)$$

where the symbol $\backsim$ means "equivalent to" in the above sense.

So, when equivalence is defined by means of a cost function J, the identification problem is reduced to an <u>optimization problem</u>: find a model $M \in C_\alpha$ such that the cost function J is minimum.

In such case, the questions related to the existence and uniqueness of the solution are the main problems of system identification. Some studies of these problems have been done by Bellman and Åström [2] for a simple class of linear systems.

The class C_α is called <u>identifiable</u> if the optimization problem has a unique solution.

At this stage some diagrams could be helpful to one visualizes the identification procedure.

Let Y be the space of outputs, such that:

$$y_S = y_S(w) \in Y \quad ; \quad w \in W$$
$$y_M = y_M(w) \in Y \quad ; \quad w \in W, \quad M \in C_\alpha$$

Assume Y is a metric space[4] and define the cost function J as the metric d in Y. That is:

$$J = d(y_M, y_S)$$

So, the identification problem (see fig. 1) can be formulated as follows: find a model $M \in C_\alpha$ such the distance from its output y_M to the system output y_S, is the smallest possible.

The figure 2 shows the same situation represented by the engineering viewpoint; that is, by block diagrams (for simplicity the outputs are assumed to be scalars and no external noise is supposed to corrupt them).

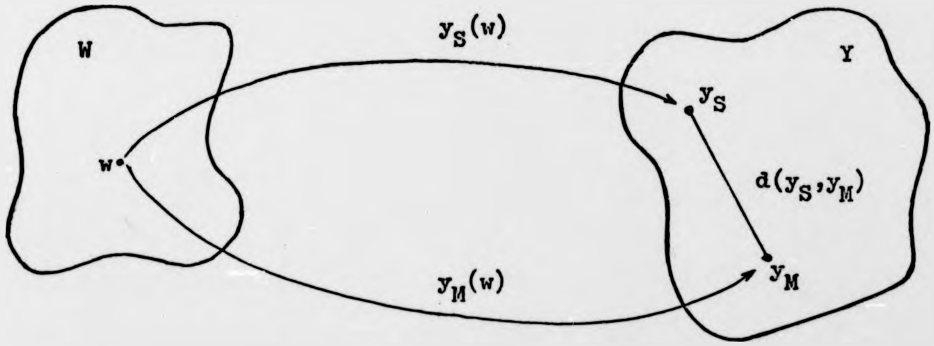

Fig. 1: $S \sim M^* \in C_\alpha \iff d(y_S, y_{M^*}) = \underset{M \in C_\alpha}{\text{Min.}}\ d(y_S, y_M)$

---

4 Actually, it would be necessary some further requirements (as linearity, inner product, and completeness) in the algebraic-topological structure of the output space Y, when the identification problem is reduced to an optimization one. In this case Y becomes a Hilbert space.

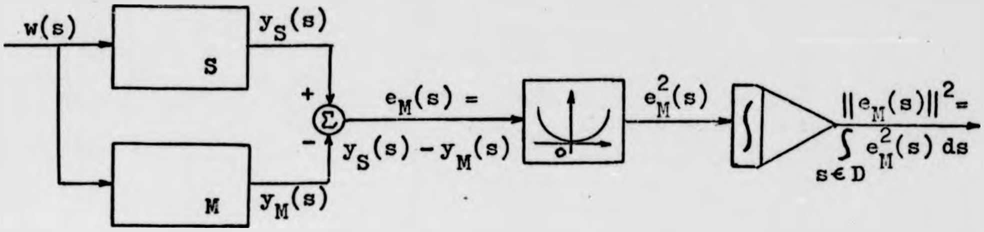

Fig. 2: $S \sim M^* \in C \iff \|e_{M^*}\| = \underset{M \in C}{\text{Min.}} \|e_M\|$

Now let us return to the essential meaning of system identification, as formulated above. Generally speaking, there are two possible ways to determine, based on the observations of input and output, the mathematical model of a given physical (or economical, or sociological, or biological, etc.) system:

1) AXIOMATIC APPROACH: Mathematico-physical (or economical, or sociological, or biological, etc.) analysis based on the laws which govern the underlying "applied" subject.

2) EXPERIMENTAL APPROACH: Data analysis where the main information about the system is obtained by measurements.

Very often the first way is used in the system classification stage. That is, we can use the mathematico-physical analysis, based on the physical laws, as some "a priori" information about the structure of a physical system S, and so a class C_α can be determined (i.e., system classification). Since a class C_α is available, a model

M in C_α may be determined by means of experimental measurements in the system S using, for example, the meaning of equivalence stated before (i.e., system identification).

It is important to notice what we mean by "a priori" knowledge (or "a priori" information). We will use this term in a very wide sense. That is, it will mean all knowledge (or information) we have about the system before we start the identification procedure.

From this viewpoint, the system classification can be considered, by itself, as "a priori" knowledge of the system's structure.

So, if the solution of the classification problem provides us with some available class of models, say C_α, all information we have about the models in C_α will be considered as "a priori". This will be the case even if the system classification was carried out by experimental analysis (some authors use the term "a priori" only for information obtained by means of non-experimental analysis).

Example 2. Assume it was determined that a given system S is represented in terms of ordinary differential equation (1st. step: System Characterization) which is linear (in the usual meaning) and of order n (2nd. step: System Classification).

Let C_n be the class of all linear ordinary differential equations of order n.

So, using "a priori" information about the system S we could classify it as of class C_n.

Each linear ordinary differential equation of order n is a model belonging to the class C_n. There are infinitely many models in C_n (actually the class C_n is uncountable), and they are completely characterized. The only difference between any two of these models is just a set of n+1 real constants, which are the coefficients of each linear ordinary differential equation in C_n. These coefficients are called the parameters of the model.

The question is: Which are these n+1 parameters? In other words, which model M (represented by its n+1 parameters) is equivalent to the system S?

1.2 - LUMPED PARAMETER SYSTEMS AND DISTRIBUTED PARAMETER SYSTEM

Before introducing the idea of "lumped parameter systems" and "distributed parameter systems", we will present an informal discussion about the meaning of the terms "dynamic systems" and "parametric models".

A system S is said to be instantaneous if it is represented in terms of a mathematical model M whose the outputs y_M at any time t depends only on the input values at the same time t. No past or future values of the input will affect the present value of the output. This may also be called a zero-memory or a memoryless system. Otherwise the system is said to be dynamic and to have memory.

If a dynamic system is one whose the model outputs do not depent on future values of the input, it is called causal (or physical, or nonanticipatory). If this is not the case the dynamic system is called noncausal (or nonphysical, or anticipatory). If a causal system is such that the model outputs depend on the past inputs only over a finite period, say T, then it is said to have finite-memory, and T is its memory length. [5]

Systems can be represented in many differents ways, as we have already seen in the last section. Now we will introduce two important disjoint classes of models for the system characterization problem: By nonparametric we mean such models described in terms of impulse response, transfer functions, covariance functions, spectral densities, etc. By parametric models we mean those ones described in terms of state equation (or more generally "dynamical equations", which means the set of equations that describes the unique ralation between the input, the output and state), differential (or difference) equations (both partial and ordinary), etc. Loosely speaking, a model is said to be parametric when it is completely characterized by a set of parameters (which can be constants, time and/or space varying, state independent, etc.). In rough terms, this means that the identification procedure is reduced to a problem of finding a certain number of parameters which completely determines the underlined model. Otherwise, it is said to be nonparametric (e.g., when the identifica-

---

5 We have avoided discussing the term dynamical system in order to keep this introductory chapter on an informal level, reducing the abstract mathematical notation to a minimum. Generally, "dynamical" and "dynamic" are slightly different concepts. In few words: "dynamical" has roughly the same meaning as "causal". For a detailed mathematical definition and interpretation of the axioms, the reader is referred to [3] - [6] .

tion procedure is reduced to the problem of finding an impulse response function belonging to some specific function space).[6]

When the system under study is represented by a parametric model, the terms <u>parametric estimation</u>[7] and <u>structure identification</u> are sometimes used to specify what we are calling "system identification" and "system classification", respectively.

Our main subject in this work will be the identification of a certain type of parametric models. So, it would seem to be a good choice to use the term "parametric estimation" instead of "system identification". But we will avoid (where possible) using the word "estimation" to specify a identification problem, reserving this term only for the problem of "state estimation". Later in this chapter, we will present some comments about the confrontation between the problems of system identification and state estimation.

Now let us return to the main topic of this section. Many authors [10] - [16] define lumped parameter systems and distributed parameter systems in some slightly different ways. Sometimes, the

---

6 The nonparametric representation has the advantage that it is not necessary to specify the order of the model explicitly. They are, however, intrinsically infinite-dimensional models. Interesting aspects of parametric versus nonparametric models can be found in the literature on time-series analysis [7] - [9] .

7 The term "parametric estimation" is also used in some wider sense, even when the models are not classified as parametric ones but the identification problem is reduced to that one of finding some unknown parameters (e.g., in the determination of a transfer function which is completely known up to a finite set of constant parameters).

properties which are used by one author as definition, are used by another one as consequence and "vice versa". For the purpose of our further studies a very brief and simple definition will be sufficient:

"A dynamic system that can be represented in terms of a ordinary differential (or difference) equation will be called a <u>lumped parameter system</u> (LPS). When it requires the use of partial differential[8] equation to describe its dynamic behavior, it will be called a <u>distributed parameter system</u>[9] (DPS)".

In brief words: LPS and DPS are characterized by finite and infinite-dimensional state space, respectively.

The meaning of the terms "lumped" and "distributed" can be better understood when the physical implications of the above definitions are more deeply analysed.

In a lumped parameter system the physical size of the system is not important, since the excitations are transmited through the system instantaneously. This assumption is usually valid if the largest physical dimension of the system is small compared to the wavelength of the highest significant frequency considered. Also, in this case, the system can be decomposed into a finite number of components, each

---

[8] The particular case of systems described by partial difference equation can be viewed as a result of an approximation method which reduces a DPS to a LPS.

[9] Dynamic systems whose mathematical models are in the form of integral (or integro-differential) equation are also called DPS.

with a finite number of input and outputs. On the other hand, in a distributed parameter system, the spatial configuration is important and generally it has dimensions that are not small compared to the shortest wavelength of interest.

Examples of physical systems that can be modeled by partial differential equations will be considered in the next chapter. Supplementary discussions concerning with applications and related topics in DPS can be found in the literature dealing with the control problem (see, for example, [10] and [17]).

1.3 - SYSTEM IDENTIFICATION VERSUS STATE ESTIMATION

First of all, let us introduce the notion of state and then the formulation of the state estimation problem.

The state notion: Actually, from a physical viewpoint, the concept of state can be thought as a primitive one, and as such it is not to be defined. Therefore some authors [18] - [26] present a "formal definition" of state which is valid, and very useful indeed, in the sense of giving a deeper insight about the meaning of state. Generally speaking, "the state u of a dynamic system at time $t = t_o$, is the amount of information at t_o that, together with any input w belonging to an input space W and known for all $t \geq t_o$, determine uniquely the "behavior" of the system". [10] But what does the term

[10] Following our previous intention of keeping this introductory chapter on an informal level, we have avoided discussing a more sophisticated but precise definition of state based on abstract mathematical concepts. The interested reader is referred to [27] - [31].

"behavior" mean in this case? It means "the state itself and the output" of the system. So, we may not use the above as a "precise definition" of state, since the intrinsic meaning of state is assumed to be known "a priori".

Remark: We talked about "state of a dynamic system". This is a slight abuse of nomenclature, since we have been using the terms "system" and "model" with different meanings. Actually the concept of state is inherent to that of "oriented abstract objects" [27] which means, in general terms, our mathematical models. Based on our previous terminology it would be more correct to say: state of a model M that characterizes a dynamic system S. Or, when we are considering the equivalent model M*, the "state of S" can be thought as the state of M*. From now on, the term "state of a dynamic system" will be used in the above sense.

Some authors use (or abuse of) the term "system identification" or even "parameter estimation" to specify a state estimation problem. Our main goal in this section will be to emphasize the difference between these two concepts.

For sake of simplicity, we will concern ourselves with the particular problem of state estimation in lumped parameter systems (finite-dimensional case), that can be modeled by a linear ordinary differential equation (i.e., the Kalman-Bucy filter [32], [33]).

Currently there are lots of books [4], [34]-[68] at many different levels dealing with the state estimation problem in finite-dimensional case.[11] The same is not true in the infinite-dimensional

---

[11] For a general review see [69].

case, specially for distributed parameter systems described by partial differential equations [70]-[73]. A recent paper by Curtain [74] gives an unified survey of this field, emphasizing the mathematical problem of rigorously modelling distributed noise. For those readers who are familiar with Kalman-Bucy filter in lumped parameter systems, the quite readable but formal works of Meditch [75], [76] are suggested as background before becoming involved with the sophisticated mathematical aspects (such as Sobolev spaces and other necessary but nontrivial concepts) which are inherent to the study of state estimation for distributed parameter systems.

Problem Formulation: Let us consider a dynamic system S modeled by a linear ordinary differential equation, whose state as a funtion of time is an n-dimensional stochastic process $\{u(t_1);\ t_1 \in T\}$, where T is some appropriate index set (an orderd subset of the reals that has a minimum element called t_0). We are interested in knowing the value of $u(t_1)$ for some fixed t_1, but $u(t_1)$ is not directly accessible to us. Suppose we can have access only to an observation process $\{z(\tau);\ t_0 \leq \tau < t,\ t \in T\}$ which is related to $u(t_1)$ by means of a linear causal system.

Let us introduce some notation:

1) Denote an estimate of $u(t_1)$ based on the measurements of the observation process $\{z(\tau)\}$ by $\hat{u}(t_1|t)$, such that

$$\hat{u}(t_1|t) = K_{t_1}[z(\tau)] \quad ; \quad t_0 \leq \tau < t \quad , \quad t \in T$$

where K_{t_1} is some linear operator defined in the observation space.

2) Let $\tilde{u}(t_1|t)$ be the estimation error which is defined by the relation

$$\tilde{u}(t_1|t) = u(t_1) - \hat{u}(t_1|t)$$

3) $L[\tilde{u}(t_1|t)]$ will be some admissible cost function of the estimation error. A typical example would be:

$$L = L[\tilde{u}(t_1|t)] = \|\tilde{u}(t_1|t)\|^2$$

where $\| \ \|$ means the usual norm in an n-dimensional Euclidean space.

4) Since $u(t_1)$ and $\hat{u}(t_1|t)$ are random vectors, it follows that $\tilde{u}(t_1|t)$ will also be a random vector and so L will be a random function. In order to get a useful measure of the error, we can define a performance criterion J as the mean value of L, that is

$$J[\tilde{u}(t_1|t)] = E\{L[\tilde{u}(t_1|t)]\}$$

where $E\{ \ \}$ stands for the espectation of a random variable in the usual meaning.

5) We say that an estimate $\hat{u}(t_1|t)$ which minimizes $J[\tilde{u}(t_1|t)]$ is a "best" or optimal estimate.

6) The linear operator K_{t_1} which give us the optimal estimate will be called the "best" or optimal linear filter.

Problem statement: Given the measures of the observation process $\{z(\tau);\ t_o \leq \tau < t;\ t \in T\}$, determine the "best" estimation $\hat{u}(t_1|t)$ of $u(t_1)$; or equivalently: determine the "best" linear filter K_{t_1}.

If $t_1 > t$, the problem is one of prediction; if $t_1 = t$, one of filtering; and if $t_1 < t$, one of smoothing or interpolation.

So we have two distinct optimizations problems namely system identification and state estimation. The system identification (or parametric estimation, as we are dealing with parametric models) is concerned with the problem of finding a set of parameters that speci-

3) $L[\tilde{u}(t_1|t)]$ will be some admissible cost function of the estimation error. A typical example would be:

$$L = L[\tilde{u}(t_1|t)] = \|\tilde{u}(t_1|t)\|^2$$

where $\| \ \|$ means the usual norm in an n-dimensional Euclidean space.

4) Since $u(t_1)$ and $\hat{u}(t_1|t)$ are random vectors, it follows that $\tilde{u}(t_1|t)$ will also be a random vector and so L will be a random function. In order to get a useful measure of the error, we can define a performance criterion J as the mean value of L, that is

$$J[\tilde{u}(t_1|t)] = E\{L[\tilde{u}(t_1|t)]\}$$

where $E\{ \ \}$ stands for the espectation of a random variable in the usual meaning.

5) We say that an estimate $\hat{u}(t_1|t)$ which minimizes $J[\tilde{u}(t_1|t)]$ is a "best" or <u>optimal estimate</u>.

6) The linear operator K_{t_1} which give us the optimal estimate will be called the "best" or <u>optimal linear filter</u>.

<u>Problem statement</u>: Given the measures of the observation process $\{ z(\tau);\ t_0 < \tau < t;\ t \in T\}$, determine the "best" estimation $\hat{u}(t_1|t)$ of $u(t_1)$; or equivalently: determine the "best" linear filter K_{t_1}.

If $t_1 > t$, the problem is one of <u>prediction</u>; if $t_1 = t$, one of <u>filtering</u>; and if $t_1 < t$, one of <u>smoothing</u> or <u>interpolation</u>.

So we have two distinct optimizations problems namely system identification and state estimation. The system identification (or parametric estimation, as we are dealing with parametric models) is concerned with the problem of finding a set of parameters that specifies a model, say M*, which is equivalent in some sense to the system S (or, which is the "best" model, based on some performance criterion, among all models M belonging to a pre-selected class of models C_{α}). Differently, the state estimation concerns with the problem of finding the "best" estimate $\hat{u}$ of the state u of a system S, which is presupposed to be fully characterized by a completely known model M*.

In recent years many researchers dedicated a great deal of their attention to the identification problem for lumped parameter systems. A large number of books [35],[49],[53],[60],[62],[64],[66],[77]-[80], surveys [81]-[91] and comparisons of different methods [92]-[98], were written about this subject. On the other hand, the bibliography on identification of distributed parameter systems is not so large. Very few books [60] deal, even superficially, with this subject; and very often the state estimation problem in systems described by partial differential equations is wrongly termed "identification" (or even, "parameter estimation"!). In the next chapter we present a survey of this field.[12]

[12] For previous surveys in DPS identification see [90] and [99].

CHAPTER 2

DISTRIBUTED PARAMETER SYSTEMS IDENTIFICATION: A SURVEY

This chapter treats the parameter identification problem in distributed systems. The various identification methods are grouped into three disjoint classes, namely: "Direct Methods", "Reduction to a Lumped Parameter System" and "Reduction to an Algebraic Equation". Under this classification we give a general survey of the main approaches to the problem of identifying distributed parameter systems.

The meaning of "parametric models", "distributed parameter systems" and "systems identification" are to be understood as introduced in the previous chapter. Standard abbreviations such as

DPS: Distributed Parameter System(s)
LPS: Lumped Parameter System(s)
ODE: Ordinary Differential Equation(s)
PDE: Partial Differential Equation(s)

will be used in this and later chapters.

2.1 - INTRODUCTION

It has become customary to begin a survey paper dealing with any problem in DPS, by remarking on two fundamental points:

1) Problems involved with DPS are much more difficult than those with LPS.

2) Very little literature has been written about DPS, compared with what has been done for LPS.

Actually these remarks still remain valid, but they now have much less significance than they had one or two decades ago. Not only are a large number of research papers being published in this field but also new techniques in Modern Mathematical System Theory have helped to lessen the gap between LPS and DPS identification methods.

Basically the main theoretical difficulty for identifying systems described by PDE, is due to the infinite dimensionality of the state space. Two approaches are normally used to face this problem; 1) Approximation of the infinite-dimensional model by a finite-dimension one, and 2) application of optimization techniques directly to the infinite-dimensional model. Recent works on optimization in abstract spaces, which contain the DPS and LPS identification as particular cases, have simplified the general concept of this later approach.

In modeling DPS, different authors assume different classes of parametric models, each one representing a particular case adapted to a specific physical system. The "best" choice for a class of models C_α, would be a sufficiently large class, such that all DPS described by PDE could be represented by models M belonging to some subclass of C_α. This ideal assumption is not usually satisfied in practice; mainly because of the great difficulty in developing an identification method for such a wide class and, at the same time being applicable in non-restrictive conditions (such as on-line identification, normal operating record, noisy observations, finite number of measurements, nonzero input, random inputs, etc.). As it will be seen in further sections, the identification methods for general models present one or more of those restrictions, either from theoretical or practical viewpoint. Discussions on mathematical description for DPS can be found in the literature dealing with the control problem (see, for example, [1] and [2]).

Another difficulty arises when we are dealing with numerical methods for distributed models. If the solution of PDE is available, it comes very often in the form of infinite series which must be (for numerical computation purposes) approximated by a finite one. If the explicit solution is not available, some approximation technique (such as finite-differences) will be required for simulation. So, sooner or later, we will be faced with approximation problems for numerical implementation of identification methods in DPS. Discussions about this topic can be found in [3]. For numerical methods in PDE see, for example, [4]-[14] (also see references in chapter 3 - part II).

It is obvious that in physical applications, the DPS identification is a more complex problem than LPS identification. One of the main reasons for that is due to the impossibility of taking measurements by using an infinite number of sensors continuously located all over the spatial domain. In this way, some kind of approximation may be (and usually is) necessary when dealing with real applications.

Concerning the second remark: The literature discussed in this chapter contains over 100 entries related only with DPS identification problem, and it is not exhaustive. Although lots of recent papers in this field are continuously appearing, the number of books (even those that dedicate few sections to the subject [15]) is still very scarce.

Before reviewing the various methods for identification we discuss briefly the underlying motivations.

Physical systems that can be modeled by PDE (i.e., distributed parameter physical systems) are often encountered in engineering applications: Antennas, wave guides, propagation of electromagnetic and mechanical waves, microwave tubes, transmission lines, gas lines, many fluid flow systems, heat exgengers, heat insulating slabs, mechanical torsion bars, vibrating beams and strings, physical structures, transportation, environmmental and geological systems, chemical and nuclear reactors, nuclear plasma devices, and charged particles accelerators; are just a few examples of systems whose state variables are distributed in space.

Also the majority of industrial and technological systems are characterized by the same fact (e.g., aerospace, petroleum, power, steel, glass, cement and chemical industries; ferrous and nonferrous metallurgy; drying and evaporation machinery; rolling mills; etc.).

There are a wide range of identification problems for parametric models in the real world of distributed systems. Some examples of fundamental physical parameters appearing in DPS are listed below:

1) Electromagnetic properties (e.g., conductivity, permissivity, permeability, charge density, etc.)
2) Thermal properties (e.g., specific heat, thermal conductivity, heat transfer coefficient, etc.)
3) Gas and fluid properties (e.g., density, diffusion constant, viscosity, expansion and compressibility coefficients, etc.)
4) Material properties (e.g., elasticity modulus)

5) Chemical properties (e.g., activation energy, reaction velocity constant, etc.)

Much of classical and modern, practical and theoretical engineering has been concerned with this basic problem. The first efforts in identifying such fundamental parameters has been performed under rigorously controlled laboratory conditions (normally off-line identification assuming noiseless measurements). In this survey the major attention in given to the problem of identifying coefficients in parametric models[1] for DPS from sequential data (time series), by using identification methods wich can be performed under less restrictive environmental and computational conditions.

2.2 - CLASSIFICATION OF METHODS

In the next section we will discuss several identification procedures for DPS. Although each one of them treats the problem under different conditions, they can be grouped into three different classes.

CLASS Γ_1: (Direct Methods) Consists of those methods that use optimization techniques directly to the distributed (infinite-dimensional) model.

CLASS Γ_2: (Reduction to a LPS) Consists of those methods that reduce the DPS (described by a PDE) to a continuous or discrete-time LPS (described by ODE or difference equation).

---

1 Such as DPS driven by random inputs and observed through noisy measurements, experimental data, normal operating records, recursive on-line identification algorithms, etc.).

CLASS Γ_3: (Reduction to an Algebraic Equation) Consists of those methods that reduce the PDE to an algebraic equation. (If the time variable is involved, the class Γ_3 may be viewed as a subclass of Γ_2 when a finite time interval is discretized).

The methods in classes Γ_2 and Γ_3 are characterized by two stages (as opposed to those methods in class Γ_1 which have a "single" stage). The first stage is concerned with techniques for approximating the infinite-dimensional state space to a finite-dimensional one, and the second with techniques for parametric estimation.

It is important to note that, in the case of Γ_1, the techniques for parametric estimation are generally applied after that numerical approximations have been carried out. In this way, these techniques apply to finite-dimensional systems and not all present an infinite-dimensional analysis. Actually, from this viewpoint, the "single" stage characterizing the methods in class Γ_1 could be split into two sub-stages: The first one going up to the point where numerical approximation are used for computational purposes (in this sub-stage the methods work for infinite-dimensional spaces). The second is concerned with techniques used for parametric estimation, which are applied after that point. (In this second sub-stage, the majority of the methods do not present an infinite-dimensional treatment).

Since our classification is based on whether or not a method reduces the infinite-dimensional state space to a finite-dimensional one in order to apply known identification techniques, these two sub-stages appearing in class Γ_1 will not be emphasized.

Fig. 3 shows a diagram representing these three classes, where the paths (1), (2) and (3) correspond to Γ_1, Γ_2 and Γ_3 respectively. The bifurcation (2-A) and (2-B), appearing on path (2) characterizes the possibility of reducing the DPS to a discrete or continuous-time LPS. The link (2-3) represents the possibility of reducing to an algebraic equation via an ODE.

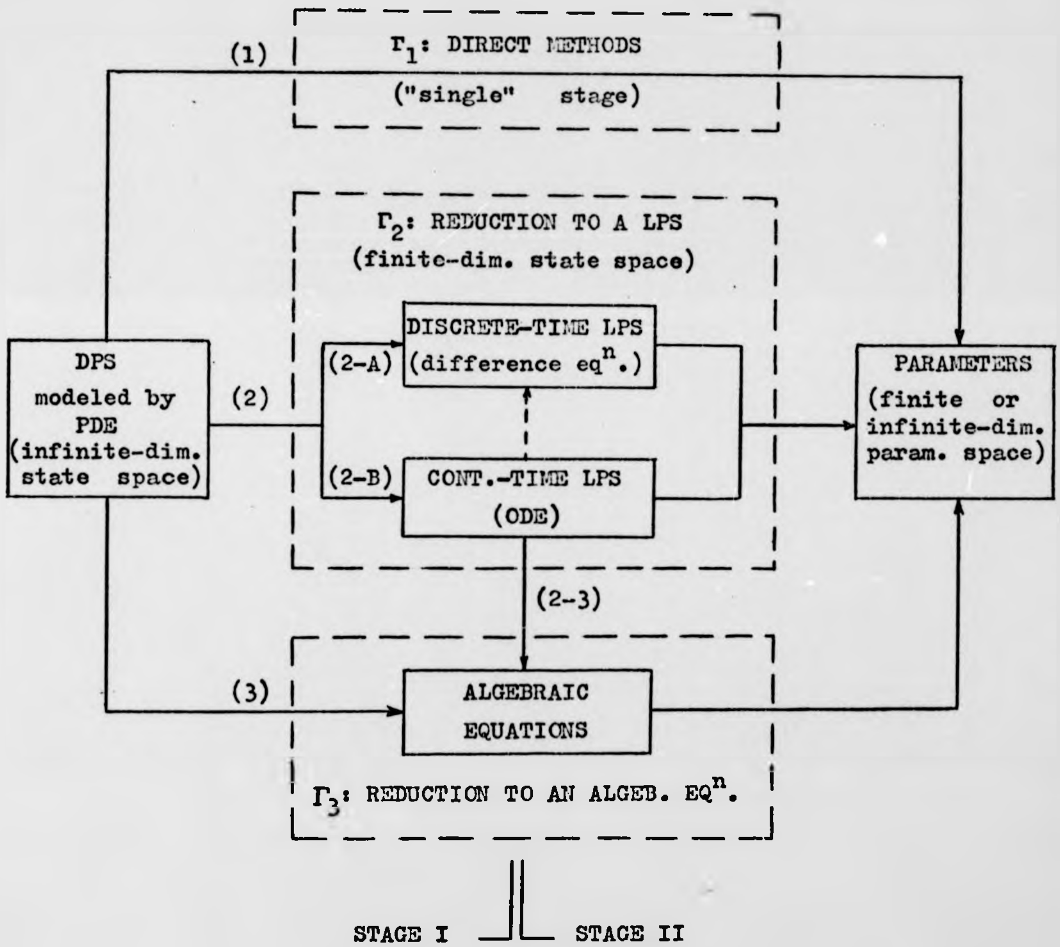

Fig. 3: Classification of the Identification Procedures for DPS.

tively. The bifurcation (2-A) and (2-B), appearing on path (2) characterizes the possibility of reducing the DPS to a discrete or continuous-time LPS. The link (2-3) represents the possibility of reducing to an algebraic equation via an ODE.

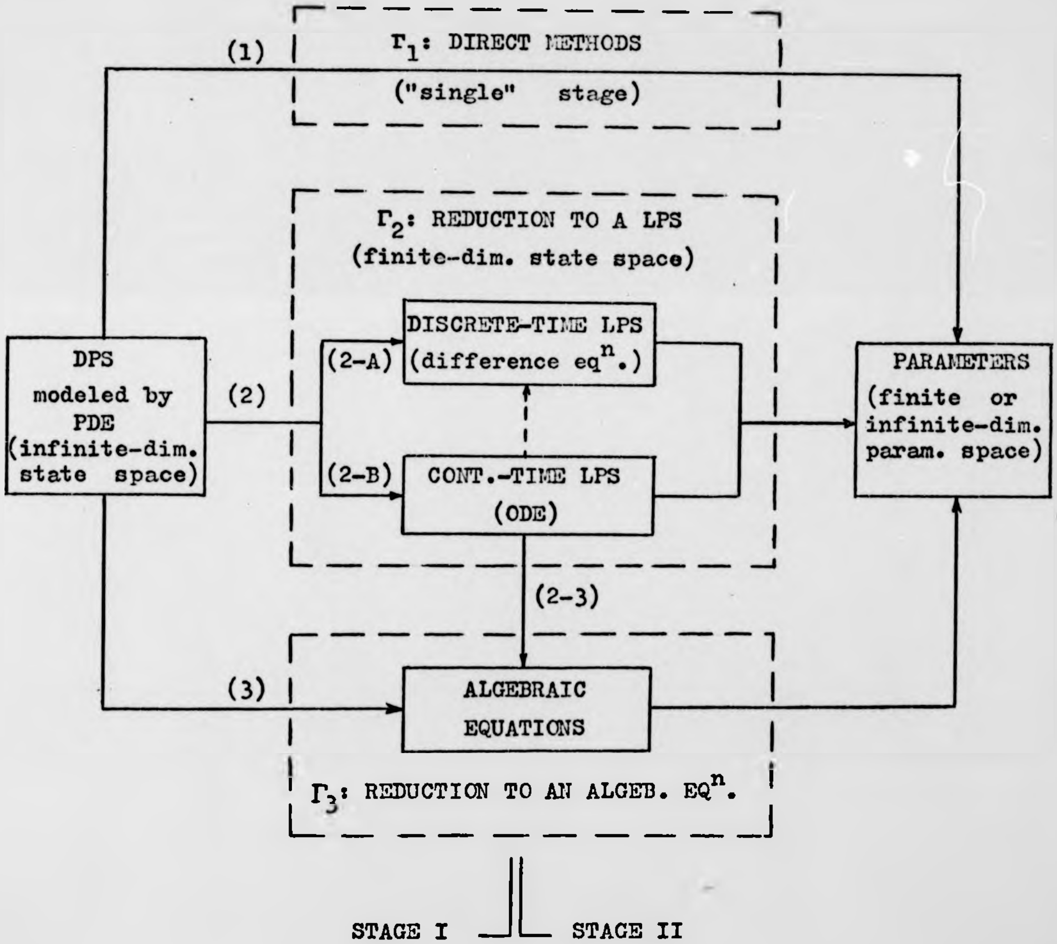

Fig. 3: Classification of the Identification Procedures for DPS.

Concerning the methods that will be discussed on section 2.3, we can classify them as follows (where the numbers between brackets correspond to the references listed at the end of this work).

<u>CLASS Γ_1</u>: Path (1): [56]-[61], [75], [80], [81], [83], [88], [91]-[104].

<u>CLASS Γ_2</u>: Path (2-A): [50]-[53], [85].
Path (2-B): [57], [65], [70]-[72], [76]-[79], [84], [87], [89], [90].

<u>CLASS Γ_3</u>: Path (3): [45], [46], [54], [55], [86].
Path (2-3): [66].

The majority of the methods belonging to Γ_2 and Γ_3 use the following techniques for stage I:

Path (2-A): Finite-differences[2]: [50]-[53], [85].

Path (2-B):
- Method of lines[3]: [57], [87].
- Method of characteristics (e.g., see [73], [74]): [70]-[72].
- Galerkin's method (e.g., see [7], [10]): [76]-[79].
- First order perturbation: [84].
- Integral transformations: [65].

Path (2-3): Method of line + Integral transformations: [66].

Path (3):
- Finite-differences: [54], [55].
- Integral transformations: [45], [46], [86].

Finally we mention the most pertinent techniques used for parametric estimation in the DPS identification problem. These tech-

---

2 Very often finite-difference techniques for approximating partial derivatives (e.g., see [4], [8], [13]) are also used by methods of class Γ_1 for numerical implementation (e.g., see [80], [83], [97], [101]).

3 Basically, finite-difference techniques applied only over the spatial domain (e.g., see [67]-[69]).

niques are concerned with stage II for methods in Γ_2 and Γ_3, and with the "single" stage for methods in Γ_1.

1) Gradient (Conjugate Gradient - Steepest Descent): [56]-[61], [71], [76]-[79], [80], [81], [97], [100], [101].

2) Stochastic Approximation: [50]-[53], [70], [72], [85].

3) Least Squares (Sequential Weighted Least Squares - Least Squares Filtering): [58], [59], [88], [89].

4) Kalman-Filter: [84].

5) Nonlinear Filtering: [77]-[79].

6) Nonlinear Programming: [83].

7) Maximum Likelihood: [51], [103].

8) Statistical Decision Theory: [51].

9) Pattern Search: [77], [79].

10) Quasilinearisation: [57], [61], [90].

11) Methods for Solution of Algebraic Equations: Class Γ_3.

Generally, in case of methods in class Γ_1, these estimation techniques are not developed using infinite-dimensional analysis, as commented before. For an interesting study concerning infinite-dimensional gradients of functionals over an infinite-dimensional parameter space, see Chavent [100].

2.3 - A CONCISE GENERAL REVIEW

The first attempts to identify parameters in DPS were mainly due to investigations dealing with the "Inverse Problem in Heat Trans-niques are concerned with stage II for methods in Γ_2 and Γ_3, and with the "single" stage for methods in Γ_1.

1) Gradient (Conjugate Gradient - Steepest Descent): [56]-[61], [71], [76]-[79], [80], [81], [97], [100], [101].
2) Stochastic Approximation: [50]-[53], [70], [72], [85].
3) Least Squares (Sequential Weighted Least Squares - Least Squares Filtering): [58], [59], [88], [89].
4) Kalman-Filter: [84].
5) Nonlinear Filtering: [77]-[79].
6) Nonlinear Programming: [83].
7) Maximum Likelihood: [51], [103].
8) Statistical Decision Theory: [51].
9) Pattern Search: [77], [79].
10) Quasilinearisation: [57], [61], [90].
11) Methods for Solution of Algebraic Equations: Class Γ_3.

Generally, in case of methods in class Γ_1, these estimation techniques are not developed using infinite-dimensional analysis, as commented before. For an interesting study concerning infinite-dimensional gradients of functionals over an infinite-dimensional parameter space, see Chavent [100].

2.3 - A CONCISE GENERAL REVIEW

The first attempts to identify parameters in DPS were mainly due to investigations dealing with the "Inverse Problem in Heat Transfer" [16]-[27][4]. Two different approaches were initially used to attack the identification problem: 1) Methods based on the analytical solution of PDE [28]-[38],[141], and 2) methods developed in frequency domain [39]-[42] (e.g., identification of some coefficients of a transfer function[5] approximating a linear model for distributed systems).

The purpose of this section is a brief review of the more relevant literature dealing with parameter identification in distributed systems. The bibliography mentioned here has been published during the last decade, and it is widely available.

Some survey papers have already appeared in this field.

Kozhinsky and Rajbman [43] and Rajbman [44], discuss the work done in the Soviet Union. They present an extensive list of references and lots of applications. A general inspection of system identification problems (both for LPS and DPS) is also considered [44], including analysis of mathematical models accuracy [43].

A recent survey was presented by Goodson and Polis [3]. They consider a "step by step" approach[6] wich treats the identification prob-

---

4 Although the label "Inverse Problem" has originated from classical solutions of identification problems, it is still used in very recent papers [92]-[94], [103], [105], [106], where a modern abstract approach is applied to solve the old problem.

5 As observed in the last chapter, these models are classified as non-parametric ones. Since we are interested only in parametric models, such identification methods will not be discussed here.

6 This approach has been considered previously by the same authors in [77],[79]. See also [143] for a latter and concise version of [3].

lem by separating it into seven independent subproblems. As well as a collection of five suggestions for further research in this field, they also present a very interesting bibliographical analysis where the major techniques used for identifying DPS are displayed together with their respective frequencies of usage. The state estimation problem is also considered.

A large number of methods have been developed to solve the problem. The great majority of them is restricted to particular cases (such as specific classes of parametric models, boundary conditions, input signals and output measurements).

Perdreauville and Goodson [45],[46] used integration by parts (an extension of Shinbrot's technique[7]) to reduce the distributed model to a set of algebraic equations. The method is applicable to nonlinear models (where extraneous terms may be included) and the case of space-varying coefficients is also considered. Normal operating records and experimental data may be used. The main limitations of the method are: 1) It is not convenient for on-line applications, 2) each model has to be considered separately, 3) it has a restricted applicability, since it requires the choice of a suitable function (wich is not always easy or possible) to multiply the PDE in order to perform the integral transformation and, 4) no noise observations were assumed.

---

7 The "modulating function method" of reducing ODE to a set of algebraic equations proposed by Shinbrot [47], was previously utilized by Loeb and Cahen [48], and Takaya [49] for identifying parameters in LPS.

Zhivoglyadov and Kaipov [50] applied finite-difference [4], [8],[13] techniques to reduce a time-invariant DPS (whose model is not necessarily linear) to a discrete-time LPS. Estimates for constant unknown parameters were obtained by minimizing a performance criterion. In this way, they compute the gradient of a cost function for each different model. Assuming noisy observations taken at discrete points in space, a stochastic approximation algorithm is used as a searching scheme for finding these estimates. The method is suitable for on-line applications. In [51] they develop some DPS identification methods based on statistical decision theory, maximum likelihood and stochastic approximation. In [52] the accuracy of a stochastic approximation method is analysed. The work of Zhivoglyadov and his group is summarized in [53].

Collins and Khatri [54], [55], assumed a deterministic class of DPS described by a time-varying model which can be nonlinear in the dynamics, but must be linear in the parameters. Based on this linearity, the q-constant vector to be identified is placed in an explicit form, and finite-difference techniques are used to approximate the partial derivatives. This procedure reduces the identification problem to that of solving a q-dimensional linear algebraic equation, where a $q \times q$-matrix must be inverted. The observations are taken at a finite number ($>q$) of points in time and/or space, and extraneous terms may be included in the original model. Normal operating data and on-line identification may be used, but measurements are assumed to be noiseless.

Seinfeld and Chen [56]-[61] developed methods for nonlinear DPS identification based mainly on the steepest descent algorithm. In [56], systems modeled by hyperbolic or parabolic PDE with constant parameters are identified. A steepest descent algorithm is used as an optimal scheme for minimizing a quadratic error criterion, where the concept of sensitivity coefficients is introduced. Analysis of output transformations and observability for DPS[8] are developed, but no noise in the observations is assumed. The method is not convenient for on-line applications and requires integration of PDE. In [57] and [61] they used steepest descent, quasilinearization and collocation techniques for nonsequential (off-line) estimation of constant parameters. Optimal location of measurements is also considered. The identification of space-varying parameters is developed in [58] and [59], where two techniques are presented: steepest descent and least-squares filtering. In [60] they considered estimation of time and/or space-varying parameters, and also of those which govern the spatial domain. The identification problems are formulated as optimal control problems and necessary conditions for optimality are derived. The techniques of steepest descent and conjugate gradient are applied. In [57]-[61] they assume noisy observations.

Tzafestas [65] considered the estimation of constant parameters in linear stochastic DPS, which can be reduced to an equivalent LPS by means of integral transformations. Although the method is basically the same used by Perdreauville in [46] and no noisy observations have been considered, this was one of the first attempts to identify DPS driven by random inputs. He assumes a discrete-time analogue of the original model and uses integration by parts to get a canonical LPS. In [140] a particular class of discrete-time DPS is considered.

---

8 Studies concerning observability in DPS can be found in [1],[62]-[64].

rameters are identified. A steepest descent algorithm is used as an optimal scheme for minimizing a quadratic error criterion, where the concept of sensitivity coefficients is introduced. Analysis of output transformations and observability for DPS[8] are developed, but no noise in the observations is assumed. The method is not convenient for on-line applications and requires integration of PDE. In [57] and [61] they used steepest descent, quasilinearization and collocation techniques for nonsequential (off-line) estimation of constant parameters. Optimal location of measurements is also considered. The identification of space-varying parameters is developed in [58] and [59], where two techniques are presented: steepest descent and least-squares filtering. In [60] they considered estimation of time and/or space-varying parameters, and also of those which govern the spatial domain. The identification problems are formulated as optimal control problems and necessary conditions for optimality are derived. The techniques of steepest descent and conjugate gradient are applied. In [57]-[61] they assume noisy observations.

Tzafestas [65] considered the estimation of constant parameters in linear stochastic DPS, which can be reduced to an equivalent LPS by means of integral transformations. Although the method is basically the same used by Perdreauville in [46] and no noisy observations have been considered, this was one of the first attempts to identify DPS driven by random inputs. He assumes a discrete-time analogue of the original model and uses integration by parts to get a canonical LPS. In [140] a particular class of discrete-time DPS is considered.

---

[8] Studies concerning observability in DPS can be found in [1],[62]-[64].

Fairman and Shen [66] modified the method of Perdreauville and Goodson [46] of reducing the model for DPS to a set of algebraic equations. They applied the method of lines to convert PDE into ODE[9]. This way, they avoid two restrictions required in [46], namely: off-line numerical integration and complete knowledge of the state, both over the spatial domain. In a second step they reduce the ODE to a set of algebraic equations by using the "moment functional method"[10]. This procedure has been applied to identification of constants parameters in one-dimensional wave and diffusion equations. A particular case of a time-varying coefficient was also considered. The observations were taken at a finite number of points in space, but were assumed to be noiseless.

Carpenter, Wozny and Goodson [70]-[72], used the method of characteristics [73], [74] to reduce a linear first-order PDE to a set of ODE. Estimates of unknown parameters (which may depend on the independent variables and states) were obtained by minimizing a quadratic performance criterion. Stochastic approximation algorithms were chosen as a recursive searching scheme for finding the estimates. They assumed noisy observations and limited available measurement transducers, but the on-line applicability of the method depends on the required performance criterion.

---

9 The method of lines (e.g., see [67]-[69]) reduces the DPS to a continuous-time LPS (i.e., difference-differential equations).

10 Basically, this method consists in multiplying the ODE by a suitable modulating function (they used a modified form of the Poisson probability density function) and then integrating by parts.

Ruban [75] presented an algorithm for identifying DPS by means of sensitivity functions. This method is similar to that proposed by Seinfeld in [56].

Polis, Goodson and Wozny [76]-[79] proposed a step by step approach to the DPS identification problem. They assumed an approximate solution for the distributed model, based on a finite set of orthogonal functions over the spatial domain. The Galerkin's criterion [7], [10] is used to reduce the PDE to a set of ODE. The constant parameters are then identified by means of known techniques for LPS identification. Three optimization schemes were considered to minimize a performance criterion: steepest descent, nonlinear filtering and pattern search. Decision for measurement locations in the spatial domain were taken based on the G-K observability [62]. Noisy observations and extraneous terms were considered.

Di Pillo and Grippo [80] applied the "epsilon technique"[11] to estimate constant parameters and states in linear DPS. Noisy observations were taken at a finite number of points in the spatial domain, and finite-difference techniques were used for numerical implementation of the method. In [81] they proposed an alternative procedure, by using an approximate solution, to avoid finite-difference approximations.

Hamza and Sheiran [83] presented a method for identifying constant and time-varying parameters in DPS. Non-linear programming was applied to minimize a discrete version of an appropriate perfor-

---

[11] Basically, the ϵ-technique consists in minimizing a new cost function, which is obtained by adding to the original one a penalty term. For details see [82].

mance criterion, where finite-difference techniques are used to approximate the partial derivatives. In the case of linear models and assuming the instantaneous error squared as a performance criterion, the method reduces to one similar to that proposed by Collins and Khatri in [55]. This identification procedure is suitable for on-line application and uses a limited number of sensors along the spatial domain. Examples considering noisy measurements, extraneous terms and experimental results were included.

Bhagavan and Nardizzi [84] considered the identification of DPS modeled by linear PDE with constant coefficients. First-order perturbations were used to reduce the problem to one of estimation in LPS by means of Kalman filtering. The method is suitable for on-line applications and assumes a finite number of noisy observations.[12]

Several papers dealing with the identification problem in DPS were presented in IFAC symposiums and other international conferences. Some of them have already been commented on here.

Diamessis [86] used integral transformations and approximations by Chebychev polynomials to reduce a linear PDE with constant coefficients to a set of algebraic equations. The method is not suitable for on-line identification and does not assume noisy observations.

Luckinbill and Childs [87] applied the method of lines to reduce a quasilinear second-order PDE to an ODE, which is augmented by

---

12 In chapter 4 we propose a new identification method for a class of linear DPS operating in stochastic environment. The method is based on the stochastic approximation theory and a rather simplified version of it can be found in [85].

adjoining the constant parameters. A Newton-Raphson-Kantorovich expansion is used to solve the resulting linear boundary value problem. The identification procedure can operate under on-line conditions, but no noisy observations were considered.

Sherry and Shen [88] presented a method for parameter and state estimation in linear DPS, by using a sequential weighted least-squares algorithm. A finite number of noisy observations was considered, and the method is suitable for on-line applications.

Shridhar and Balatoni [89] used splines to reduce the DPS to a continuous-time LPS. A recursive least-squares procedure was applied for parametric estimation.

Chaudhuri [90] proposed an identification method for DPS based on differential approximation and quasilinearization.

A modern abstract approach for DPS identification has been considered recently by Chavent [91]-[101] and others [102]-[106].

In [97] and [101], Chavent proposed an off-line identification method for DPS described by a general class of deterministic models, where no specific probabilistic treatment for noisy observations was considered. The method is developed by functional analysis techniques and based on the Lions' [107] approach to control theory for DPS. It consists of minimizing a performance criterion (the quadratic error of output, which is nonquadratic with respect to the parameters), by using a conventional gradient technique (the steepest descent method was used). In order to compute the gradient of the performance criterion, he introduced the adjoint state (solution of adjoint state equation). The gradient is then derived as a functional of the adjoint and system states. Its computation requires the simultaneous solution of both system (given by the distributed model) and adjoint state equations. Fundamental problems in identification, such as existence, uniqueness and choice of minimization schemes, were discussed in some detail. Two types of models were considered: 1) Those with a finite number of constant parameters (finite-dimensional parameter space) and, 2) those with varying parameters as functions of independent variables, or states (infinite-dimensional parameter space). Applications were also included where discretization techniques, such as finite-differences, are applied for numerical implementation of the method; which was shown to work with a small number of measurements in space and/or time. In [100] he considered the identification of DPS modeled by parabolic PDE with space and state-varying parameters. Measurements are taken by a finite number of sensors located in the spatial domain. These sensors supply a mean value of the output over a small neighbourhood for each observation point. In this way, perturbations of measurements were considered, but a stochastic modeling for noisy observations is still lacking. As in [97], the method presents a rigorous mathematical treatment for minimizing a least square error criterion. An infinite-dimensional gradient of the performance criterion with respect to the parameters was defined, and expressed in terms of the system and adjoint states. The optimization algorithm was the steepest descent, and a detailed discussion on the uniqueness problem was also included.

Balakrishnan [102]-[104] considered the system identification problem (in particular, DPS) from a stochastic viewpoint. In [103] he presented a general abstract approach for identifying a class of linear DPS, previously considered by Lavrentiev [105] and Marchuk [106].

A stochastic formulation in Hilbert space was proposed, where additive white Gaussian noises are assumed to corrupt both input and observation process. The theoretical development was based on the semigroup theory of linear operators (as opposed to that proposed by Chavent in [97] and [100]), for time invariant systems operating under continuous-time assumption. The infinitesimal generator of a strongly continuous semigroup appearing in the model, was supposed to be dependent on unknown parameters. In order to obtain asymptotically unbiased and consistent estimates of those parameters, the "a posteriori" maximum likelihood technique was utilized. In [102] he considered the identification problem for both LPS and DPS, operating in a stochastic environment. The DPS case is just slightly mentioned, and shown to be included in the general model.

Other types of "identification" problem for DPS have appeared in the literature. Jones and Douglas [29], [30] identified a time-varying coefficient in a parabolic PDE. Since this coefficient necessarily appears in the boundary conditions and the identification method takes measurements only at the boundary, this problem is reduced to that of identifying boundary conditions. Ward and Goodson [108], [109] and Alvarado and Mukundan [139] also investigated the identification of boundary conditions. Wozny, Carpenter and Stein [110] presented a method for identifying Green's functions of DPS. Cannon [111] and Ikeda, Miyamoto and Sawaragi [142] considered the determination of unknown sources for a class of PDE. Saridis and Badavas [112], [113] identified solutions in DPS, but this was really a state estimation problem. The term "identification" was also used by Phillipson [114], [115] but again this was a state estimation problem.

Some practical subjects have stimulated several researches in this field. Problems related to physical structures, heating processes, transportation, economy, ecology, geology, chemical and environmental sciences; are just a few examples where applications and experimental results in identification of DPS have already been developed (see, for example, [16]-[26], [38], [39], [61], [97], [101], [116]-[136]).

Several conclusions can be drawn from the topics covered by this chapter. Some of them are presented in the next section. The classification introduced on section 2.2 can give us a general idea about the main techniques currently used for solving the identification problem in DPS. For a further collection of observations and comments concerning both with system identification and state estimation in DPS, the reader is referred to a previous survey by Goodson and Polis [3], [143].

2.4 - CONCLUSIONS

From what was discussed here we can select some basic points which deserve to be emphasized.

1) The reduction to finite-dimensional state space seems to be the most popular method used in the DPS identification problem (Methods of class Γ_2: the distributed model is approximated by a lumped one before any optimization is carried out).
2) Among the approximation techniques, finite-differences is one of the most used. Other techniques for dealing with distributed models, such as finite element method (e.g., [137]), should be investigated for identification purposes.

3) As remarked before, sooner or later, we are faced with some approximation problem (either for reducing to a finite-dimensional state space, for numerical implementation, or for physical applications). The question of when to use approximation techniques, before or after the optimization, has no final answer yet. Athans [138] suggests that any approximation should be applied as late as possible in order to retain the distributed nature of the model, until numerical results are required.

4) Among the optimization techniques for parametric estimation the gradient method is the most popular. Its "probabilistic version"[13], stochastic approximation, have also been successfully used by several authors.

5) No general method for a large class of models operating in non-restrictive conditions has been developed, and only a few authors consider the case of stochastic environment (random inputs and noisy observations).

6) There has been a lot of recent literature in this field, but it is still difficult to make any comparisons, because of the different models considered.

7) Although other survey papers dealing with DPS identification have already appeared [3],[44],[143], this seems to be the first attempt to inspect the more relevant techniques in this field without confusing two different problems, namely: system identification, and state estimation.

---

13 From a particular viewpoint, the stochastic approximation algorithms may be thought as a "probabilistic version" of the gradient method.

CHAPTER 3

MATHEMATICAL PRELIMINARIES

PART I: MODELS FOR DPS

We have previously defined a DPS as a dynamic system that can be modeled by a PDE. We begin this chapter by presenting a formal description of such models, emphasizing a special subclass of linear models which will be considered in chapter 4 for identification purposes.

3.1 - ON A GENERAL CLASS

A general class of models for DPS described by PDE can be formally written as follows:

Dynamic equation: $L(u,\underline{a},s) = f(s) + w(s)$ (1)

Boundary conditions: $L_\Gamma(u)\big|_{x=x'} = f_\Gamma(x',t) + w_\Gamma(x',t)$

Initial conditions: $L_o(u)\big|_{t=o} = f_o(x) + w_o(x)$

where:

i) $x \in X$, a simply connected open set in R^n: the spatial domain.

ii) $x' \in \Gamma$, the boundary of X.

iii) $t \in \mathcal{T}$, an interval which can be $(0,T]$, or $(0,\infty)$: the time domain .

iv) $s = (x,t) \in \Omega = X \times \mathcal{T}$.

v) $u(s)$ is a real-valued function on Ω belonging to an appropriate function space $H(\Omega)$: the dependent variable.

vi) $\underline{a}(s)$ is a vector with a finite number of components which are real-valued functions on Ω belonging to an appropriate space: the parameter vector.[1]

vii) $f(s)$, $f_\Gamma(x',t)$ and $f_0(x)$ are real-valued functions on Ω, $\Gamma \times \mathcal{T}$ and X, belonging to appropriate function spaces: the input, boundary and initial functions, respectively.

viii) $w(s)$, $w_\Gamma(x',t)$ and $w_0(x)$ are real-valued random fields[2] $\{w(s);\ s \in \Omega \subset R^{n+1}\}$, $\{w_\Gamma(x',t);\ (x',t) \in \Gamma \times \mathcal{T} \subset R^{n+1}\}$ and $\{w_0(s);\ x \in X \subset R^n\}$: disturbance processes corrupting the input, boundary and initial functions, respectively.

ix) L, L_Γ and L_0 are partial differential operators. L_Γ and L_0 are concerned with the boundary and initial conditions, and L represents a parametric distributed model.

A brief word about what we mean by "appropriate" space: The appropriate space associated with the parameter vector $\underline{a}(s)$ (i.e, the parameter space) will become clear later in the chapter 4 where the identi-

---

1 A more general case can be considered, where the parameter vector $\underline{a}(u,s)$ also depends on the dependent variable $u(s)$.

2 The term "random field" is used to denote a collection of random variables indexed by points taking values in a subset of R^n, as a natural extension of the concept of stochastic processes [1].

iii) $t \in \mathcal{T}$, an interval which can be $(0,T]$, or $(0,\infty)$: the time domain .

iv) $s = (x,t) \in \Omega = X \times \mathcal{T}$.

v) $u(s)$ is a real-valued function on Ω belonging to an appropriate function space $H(\Omega)$: the dependent variable.

vi) $\underline{a}(s)$ is a vector with a finite number of components which are real-valued functions on Ω belonging to an appropriate space: the parameter vector.[1]

vii) $f(s)$, $f_\Gamma(x',t)$ and $f_o(x)$ are real-valued functions on Ω, $\Gamma \times \mathcal{T}$ and X, belonging to appropriate function spaces:the input, boundary and initial functions, respectively.

viii) $w(s)$, $w_\Gamma(x',t)$ and $w_o(x)$ are real-valued random fields[2] $\{w(s);\ s \in \Omega \subset R^{n+1}\}$, $\{w_\Gamma(x',t);\ (x',t) \in \Gamma \times \mathcal{T} \subset R^{n+1}\}$ and $\{w_o(s);\ x \in X \subset R^n\}$: disturbance processes corrupting the input, boundary and initial functions, respectively.

ix) L, L_Γ and L_o are partial differential operators. L_Γ and L_o are concerned with the boundary and initial conditions, and L represents a parametric distributed model.

A brief word about what we mean by "appropriate" space: The appropriate space associated with the parameter vector $\underline{a}(s)$ (i.e, the parameter space) will become clear later in the chapter 4 where the identi-

---

1 A more general case can be considered, where the parameter vector $\underline{a}(u,s)$ also depends on the dependent variable $u(s)$.

2 The term "random field" is used to denote a collection of random variables indexed by points taking values in a subset of R^n, as a natural extension of the concept of stochastic processes [1].

fication problem is formulated. Of course, the structure of this space depends on the particular approach used for a specific identification problem. On the other hand, an appropriate space containing the independent variable u (i.e., the state space) is a function (or distribution) space which comes to be suitable to deal with a specific problem in PDE (e.g., some Sobolev space $H(\Omega)$ [3]). Considering our particular approach to the DPS identification problem (see chapter 4), there will be no need to go into details related with $H(\Omega)$; and there are two basic reasons for that:

1) We assume the existence and uniqueness of the solution for a given DPS modeled by a particular PDE.
2) We consider (for identification purposes) an approximated finite-dimensional version for modeling DPS.

3.2 - LINEAR MODELS

A class of linear models of M th order can be written, most generally, as follows:

$$L(u,\underline{a},s) = \sum_{i \in I^{n+1}} a_i(s) \frac{\partial^m}{\partial t^{i_0} \partial x_1^{i_1} \ldots \partial x_n^{i_n}} u(s)$$

---

3 Roughly speaking: $H(\Omega)$ is a Banach space of functions on Ω, equiped with a suitable norm, such that all partial derivatives of u up to the highest order involved in the model L are in $L^p(\Omega)$. If $p = 2$ $H(\Omega)$ is a Hilbert space. For details see, for example, [2] and [3].

fication problem is formulated. Of course, the structure of this space depends on the particular approach used for a specific identification problem. On the other hand, an appropriate space containing the independent variable u (i.e., the state space) is a function (or distribution) space which comes to be suitable to deal with a specific problem in PDE (e.g., some Sobolev space $H(\Omega)$ [3]). Considering our particular approach to the DPS identification problem (see chapter 4), there will be no need to go into details related with $H(\Omega)$; and there are two basic reasons for that:

1) We assume the existence and uniqueness of the solution for a given DPS modeled by a particular PDE.
2) We consider (for identification purposes) an approximated finite-dimensional version for modeling DPS.

3.2 - LINEAR MODELS

A class of linear models of M th order can be written, most generally, as follows:

$$L(u,\underline{a},s) = \sum_{i \in I^{n+1}} a_i(s) \frac{\partial^m}{\partial t^{i_0} \partial x_1^{i_1} \cdots \partial x_n^{i_n}} u(s)$$

---

[3] Roughly speaking: $H(\Omega)$ is a Banach space of functions on Ω, equiped with a suitable norm, such that all partial derivatives of u up to the highest order involved in the model L are in $L^p(\Omega)$. If $p = 2$ $H(\Omega)$ is a Hilbert space. For details see, for example, [2] and [3].

where:

i) M is a finite integer.

ii) $I = \{0,1,\ldots,M\}$.

iii) $i_j \in I$; $\forall\ j=0,1,\ldots,n$.

iv) $i = (i_0, i_1, \ldots, i_n) \in I^{n+1}$: an $(n+1)$ - tuple index.

v) $m = \sum_{j=0}^{n} i_j$ and such that $0 \leq m \leq M$.

So, in this case, the dynamic equation (1) can be written in the form

$$a_{i*}(s) \frac{\partial^M}{\partial t^M} u(s) = - \sum_{\substack{i \in I^{n+1} \\ i \neq i*}} a_i(s) \frac{\partial^m}{\partial t^{i_0} \partial x_1^{i_1} \ldots \partial x_n^{i_n}} u(s) + f(s) + w(s)$$

where $i* = (M,0,\ldots,0) \in I^{n+1}$; or equivalently (assuming $a_{i*}(s) \neq 0$, $\forall\, s \in \Omega$) in a state representation:

$$\frac{\partial}{\partial t} \underline{u}(s) = A(s)\, \underline{u}(s) + \underline{b}(f(s) + w(s))$$

where $A(s)$ is a $M \times M$-matrix of linear spatial-differential operators whose parameters may depend on $s=(x,t)$, $\underline{b}$ is a vector in R^M (e.g., $\underline{b} = (0,\ldots,0,1)$) and $\underline{u}(s)$ is a state vector with M components.

Remark: We have defined $\underline{a}(s)$ and $\underline{u}(s)$ as "finite-size" vectors of functions representing parameters and state, respectively. It is important to note that:

1) The parameter space may be infinite-dimensional since any component $a_i(s)$ of $\underline{a}(s)$ can be a function of x for each time $t \in \mathcal{T}$.

2) The state space is obviously an infinite-dimensional one, since for each time $t \in \mathcal{T}$, the components of $\underline{u}(s)$ are certainly functions of x.

A particular case: In chapter 4 we will be particularly interested in the following special subclass of linear parametric models: [4]

$$L(u,\underline{a},x) = L_t^N\, u(x,t) - L_x^M\, u(x,t) \;; \qquad x \in X \subset R \tag{2}$$

where L_t^N and L_x^M are linear partial differential operators such that

$$L_t^N\, u(x,t) = \sum_{m=0}^{N} \gamma_m(x)\, \frac{\partial^m}{\partial t^m}\, u(x,t)$$

$$L_x^M\, u(x,t) = \sum_{m=1}^{M} \alpha_m(x)\, \frac{\partial^m}{\partial x^m}\, u(x,t)$$

and the time-invariant (M+N+1)-dimensional parameter vector is given by

$$\underline{a} = \underline{a}(x) = (\underline{\alpha}(x),\underline{\gamma}(x)) \qquad \in R^{M+N+1}$$

with

$$\underline{\alpha}(x) = (\alpha_1(x),\ldots,\alpha_M(x)) \qquad \in R^M$$

$$\underline{\gamma}(x) = (\gamma_0(x),\ldots,\gamma_N(x)) \qquad \in R^{N+1}$$

---

4 For discussions (both from theoretical and practical viewpoints) on DPS whose models are included in such subclass, see for example [3]-[6].

PART II: HIGHER ORDER FINITE-DIFFERENCES

The classical finite-difference method is a well known technique used to obtain approximations for partial differential equations, mainly for second-order equations (e.g., see [6]-[19]).

The main goal of this second part of chapter 3 is to introduce a brief discussion of finite-difference techniques for approximating higher order partial derivatives. No attempt will be made to present a rigorous treatment dealing with specific topics such as stability, errors, and other technical aspects concerning with numerical analysis of finite-difference techniques; since this subject is widely available in the current literature.[5]

The results obtained here will be used in the next chapter for reducing DPS to LPS.

3.3 - SHIFT, DIFFERENCE AND SLOPE OPERATORS

Notation:

1) The set of all nonnegative integers, of all positive integers, of all even integers (including zero) and of all odd integers are denoted by Z, Z_+, Z_e and Z_o, respectively:

---

5 The interested reader is referred to the bibliography concerning with this second part of chapter 3, which is listed at the end of this work.

$$Z = \{0,1,2,\dots\}$$

$$Z_+ = \{1,2,3,\dots\}$$

$$Z_e = \{0,2,4,\dots\}$$

$$Z_o = \{1,3,5,\dots\}$$

As usual the three dots (...) indicate a presumed understanding about what is omited.

2) The symbols ($\bar{\ }$) and ($\tilde{\ }$) denote integer valued functions defined on Z as follows:

$$\bar{m} = \begin{cases} \dfrac{m}{2} \; ; & \text{if } m \in Z_e \\ \dfrac{m+1}{2} \; ; & \text{if } m \in Z_o \end{cases}$$

$$\tilde{m} = \begin{cases} \dfrac{m}{2} \; ; & \text{if } m \in Z_e \\ \dfrac{m-1}{2} \; ; & \text{if } m \in Z_o \end{cases}$$

Those functions will be extensively used in the remainder of this work and so we recall some of their main properties:

i) $\bar{m} + \tilde{m} = m$

ii) $\overline{m-1} = \tilde{m} \quad ; \quad \widetilde{m+1} = \bar{m}$

iii) $\overline{m+1} = \tilde{m} + 1 \; ; \quad \widetilde{m-1} = \bar{m} - 1$

iv) $\bar{m} - \tilde{m} = 2\bar{m} - m = m - 2\tilde{m} = \begin{cases} 0 \; ; & \text{if } m \in Z_e \\ 1 \; ; & \text{if } m \in Z_o \end{cases}$

<u>Definition (D - 3.1)</u>: Let B(R) be the linear space of all bounded real-valued functions on R. We define the "p th-order delta shift operator"

$$S_\delta^p : B(R) \rightarrow B(R)$$

where δ is a fixed positive real constant, as follows:

$$S_\delta^p f(x) = f(x+p\delta)$$

for any $f \in B(R)$ and $p \in R$.[6]

Remarks:

1) For any $p \in R$, S_δ^p is a linear operator on $B(R)$ under the usual definition of addition and scalar multiplication.

2) Any finite set of operators $\{S_\delta^{p_i}\}$ is linearly independent. That is, for any finite set of scalars $\{\alpha_i\}$

$$\sum_i \alpha_i S_\delta^{p_i} = 0 \leftrightarrow \alpha_i = 0 \text{ for all } i.$$

3) S_δ^p is invertible for any $p \in R$.

$$S_\delta^p S_\delta^{-p} = S_\delta^{-p} S_\delta^p = S_\delta^0 = I$$

where I stands for the identity operator on $B(R)$.

Three special linear operators on $B(R)$ are derived from S_δ^p as follows:

Definition (D - 3.2): The "forward difference operator":

---

6 With $\Theta = -p\delta$ we get $S_\delta^p = S_\Theta$, the "delay" operator: $S_\Theta f(x) = f(x-\Theta)$. For further details on shift or delay operators see, for example, [20]. In the finite-differences literature the symbol E is used instead of S to denote shift (also called "displacement") operators.

$$\Delta_\delta = S_\delta^1 - S_\delta^0 = S_\delta - I \ .$$

Definition (D - 3.3): The "backward difference operator":

$$\Delta_\delta^{-1} = S_\delta^0 - S_\delta^{-1} = I - S_\delta^{-1} \ .$$

Definition (D - 3.4): The "m+1 th-order centered slope operator":

$$D_\delta^{m+1} = \frac{1}{\delta} \Delta_\delta^{(-1)^m} D_\delta^m \ ; \quad D_\delta^0 = I \ .$$

Remark: Δ_δ^{-1} is not the inverse of Δ_δ since $\Delta_\delta \Delta_\delta^{-1} = \Delta_\delta^{-1}\Delta_\delta \neq I$.

These operators, mainly S_δ^p and D_δ^m, will be used for approximating higher order derivatives in the next section. But before going through that we need to prove the following results:

Proposition (P - 3.1):

$$(S_\delta^p + a S_\delta^q)^m = \sum_{i=0}^m a^{m-i} \binom{m}{i} S_\delta^{mq + i(p-q)}$$

where p, q and a are reals ($a \neq 0$), $m \in Z$ and

$$\binom{m}{i} = m! / [\, i!\,(m-i)!\,] \ .$$

Proof:

Both cases, m=0 and m=1, are automatically satisfied by simple substitution. Now assume the following equivalent proposition:

$$(a\, S_\delta^q)^{-m} \, (S_\delta^p + a S_\delta^q)^m = \sum_{i=0}^m a^{-i} \binom{m}{i} S_\delta^{i(p-q)} \ . \qquad (3)$$

Then

$$(a S_\delta^q)^{-(m+1)} \; (S_\delta^p + a S_\delta^q)^{m+1} =$$

$$= (a S_\delta^q)^{-m} \; (S_\delta^p + a S_\delta^q)^m \; a^{-1} S_\delta^{-q} \, (S_\delta^p + a S_\delta^q) =$$

$$= \sum_{i=0}^{m} a^{-i} \binom{m}{i} S_\delta^{i(p-q)} \; (a^{-1} S_\delta^{p-q} + I) =$$

$$= \sum_{i=0}^{m} a^{-(i+1)} \binom{m}{i} S_\delta^{(i+1)(p-q)} + \sum_{i=0}^{m} a^{-i} \binom{m}{i} S_\delta^{i(p-q)}$$

$$= \sum_{i=1}^{m+1} a^{-i} \binom{m}{i-1} S_\delta^{i(p-q)} + \sum_{i=0}^{m} a^{-1} \binom{m}{i} S_\delta^{i(p-q)}$$

$$= a^0 \binom{m}{0} S_\delta^0 + \sum_{i=1}^{m} a^{-i} \left[\binom{m}{i-1} + \binom{m}{i}\right] S_\delta^{i(p-q)} + a^{-(m+1)} \binom{m}{m} S_\delta^{(m+1)(p-q)} .$$

Since

$$\binom{m}{i-1} + \binom{m}{i} = \binom{m+1}{i}$$

$$\binom{m}{0} = \binom{m+1}{0} = 1 \quad ; \quad \binom{m}{m} = \binom{m+1}{m+1} = 1$$

we get:

$$(a S_\delta^q)^{-(m+1)} \; (S_\delta^p + a S_\delta^q)^{m+1} =$$

$$= a^0 \binom{m+1}{0} S_\delta^0 + \sum_{i=1}^{m} a^{-i} \binom{m+1}{i} S_\delta^{i(p-q)} + a^{-(m+1)} \binom{m+1}{m+1} S_\delta^{(m+1)(p-q)} =$$

$$= \sum_{i=0}^{m+1} a^{-1} \binom{m+1}{i} S_\delta^{i(p+q)}$$

which proves, by induction, the assumption (3). □

<u>Particular cases:</u>

1) $a = -1$, $p = 1$ and $q = 0$:

$$\Delta_\delta^m = (S_\delta - I)^m = \sum_{i=0}^{m} (-1)^{m-i}\binom{m}{i} S_\delta^i \quad .$$

2) $a = -1$, $p = 0$ and $q = -1$

$$\Delta_\delta^{-m} = (I - S_\delta^{-1})^m = \sum_{i=0}^{m} (-1)^{m-i}\binom{m}{i} S_\delta^{i-m} \quad .$$

Proposition (P - 3.2):

$$D_\delta^m = \begin{cases} I & ; \quad \text{if} \quad m = 0 \\ \frac{1}{\delta^m}\prod_{i=1}^{m}\Delta_\delta^{(-1)^{i-1}} & ; \quad \text{if} \quad m \in Z_+ \end{cases}$$

where

$$\prod_{i=1}^{m}\Delta_\delta^{(-1)^{i-1}} = \Delta_\delta^{(-1)^0}\Delta_\delta^{(-1)^1} \cdots\cdots \Delta_\delta^{(-1)^{m-2}}\Delta_\delta^{(-1)^{m-1}} =$$

$$= \Delta_\delta^{(-1)^{m-1}}\Delta_\delta^{(-1)^{m-2}} \cdots\cdots \Delta_\delta^{-1}\ \Delta_\delta$$

Proof:

a) $m = 0$ and $m = 1$, trivial by (D - 3.4).

b) $m > 1$: Assume

$$D^m = \frac{1}{\delta^m}\prod_{i=1}^{m}\Delta_\delta^{(-1)^{i-1}} .$$

So, by (D - 3.4), we get:

$$D^{m+1} = \frac{1}{\delta}\,\Delta_\delta^{(-1)^m}\,\frac{1}{\delta^m}\prod_{i=1}^{m}\Delta_\delta^{(-1)^{i-1}} =$$

$$= \frac{1}{\delta^{m+1}}\prod_{i=1}^{m+1}\Delta_\delta^{(-1)^{i-1}}$$

and the proof (by induction) is completed. □

Lemma (L - 3.1): The mth-order centered slope operator can be written in terms of a finite series of delta shift operators as follows:

$$D_\delta^m = \frac{1}{\delta^m} \sum_{i=0}^{m} (-1)^{m-i} \binom{m}{i} S_\delta^{i-\tilde{m}} \; ; \qquad m \in Z$$

and this representation is unique.

Proof:

Existence:

a) $m = 0$: $D_\delta^0 = S_\delta^0 = I$ (trivial).

b) $m \in Z_+$: By (P - 3.2) we have

$$\delta^m D_\delta^m = \prod_{i=1}^{m} \Delta_\delta^{(-1)^{i-1}} = \Delta_\delta \Delta_\delta^{-1} \cdots \Delta_\delta^{(-1)^{m-1}} \Delta_\delta^{(-1)^{m-2}} =$$

$$= \begin{cases} \Delta_\delta \Delta_\delta^{-1} \cdots \Delta_\delta^{-1} \Delta_\delta & ; \text{ if } m \in Z_e \\ \Delta_\delta \Delta_\delta^{-1} \cdots \Delta_\delta^{-1} \Delta_\delta \Delta_\delta^{-1} & ; \text{ if } m \in Z_o \end{cases}$$

$$= \begin{cases} (\Delta_\delta \Delta_\delta^{-1})^{\frac{m}{2}} & ; \text{ if } m \in Z_e \\ (\Delta_\delta \Delta_\delta^{-1})^{\frac{m-1}{2}} & ; \text{ if } m \in Z_o \end{cases}$$

$$= (\Delta_\delta \Delta_\delta^{-1})^{\tilde{m}} \begin{cases} I & ; \text{ if } m \in Z_e \\ \Delta_\delta & ; \text{ if } m \in Z_o \end{cases}$$

But

$$\Delta_\delta \Delta_\delta^{-1} = (S_\delta - I)(S_\delta - S_\delta^{-1}) = S_\delta - 2I + S_\delta^{-1} = (S_\delta^{\frac{1}{2}} - S_\delta^{-\frac{1}{2}})^2 ,$$

then by (P - 3.1) with $a = -1$ and $p = -q = \frac{1}{2}$ we get

$$(\Delta_\delta \Delta_\delta^{-1})^{\tilde{m}} = (S_\delta^{\frac{1}{2}} - S_\delta^{-\frac{1}{2}})^{2\tilde{m}} = \sum_{i=0}^{2\tilde{m}} (-1)^{2\tilde{m}-i} \binom{2\tilde{m}}{i} S_\delta^{-\tilde{m}+i} .$$

Hence:

$$\delta^m D_\delta^m = \sum_{i=0}^{2\tilde{m}} (-1)^{2\tilde{m}-i} \binom{2\tilde{m}}{i} S_\delta^{i-\tilde{m}} \begin{cases} I & ; \text{ if } m \in Z_e \\ \Delta_\delta & ; \text{ if } m \in Z_o \end{cases} .$$

b-i) $m \in Z_e \rightarrow 2\tilde{m} = m$ and the existence is proved for $m \in Z_e$.

b-ii) $m \in Z_o \rightarrow 2\tilde{m} = m-1$ and $(-1)^{2\tilde{m}-i} = (-1)^{m-1-i} = (-1)^{m+1-i}$. So:

$$\delta^m D_\delta^m = \sum_{i=0}^{m-1} (-1)^{m+1-i} \binom{m-1}{i} S_\delta^{i-\tilde{m}} (S_\delta - I) =$$

$$= \sum_{i=0}^{m-1} (-1)^{m+1-i} \binom{m-1}{i} S_\delta^{i+1-\tilde{m}} - \sum_{i=0}^{m-1} (-1)^{m+1-i} \binom{m-1}{i} S_\delta^{i-\tilde{m}} =$$

$$= \sum_{i=1}^{m} (-1)^{m-i} \binom{m-1}{i-1} S_\delta^{i-\tilde{m}} - \sum_{i=0}^{m-1} (-1)^{m+1-i} \binom{m-1}{i} S_\delta^{i-\tilde{m}} =$$

$$= -\binom{m-1}{0} S_\delta^{-\tilde{m}} + \sum_{i=1}^{m} \left[(-1)^{m-i} \binom{m-1}{i-1} - (-1)^{m+1-i} \binom{m-1}{i}\right] S_\delta^{i-\tilde{m}} + \binom{m-1}{m-1} S_\delta^{m-\tilde{m}} =$$

$$= -\binom{m}{0} S_\delta^{-\tilde{m}} + \sum_{i=1}^{m} (-1)^{m-i} \left[\binom{m-1}{i-1} + \binom{m-1}{i}\right] S_\delta^{i-\tilde{m}} + \binom{m}{m} S_\delta^{m-\tilde{m}} =$$

$$= (-1)^m \binom{m}{0} S_\delta^{-\tilde{m}} + \sum_{i=1}^{m} (-1)^{m-i} \binom{m}{i} S_\delta^{i-\tilde{m}} + \binom{m}{m} S_\delta^{m-\tilde{m}} =$$

$$= \sum_{i=0}^{m} (-1)^{m-i} \binom{m}{i} S_\delta^{i-\tilde{m}}$$

and the existence is proved for $m \in Z_o$.

Uniqueness:

Set

$$\alpha_i = (-1)^{m-i} \binom{m}{i} \frac{1}{\delta^m}$$

such that

$$D_\delta^m = \sum_{i=0}^{m} \alpha_i \, S_\delta^{i-\tilde{m}} \quad .$$

Now suppose also that

$$D_\delta^m = \sum_{i=0}^{m} \beta_i \, S_\delta^{i-\tilde{m}} \quad .$$

Then

$$0 = D_\delta^m - D_\delta^m = \sum_{i=0}^{m} (\alpha_i - \beta_i) \, S_\delta^{i-\tilde{m}}$$

by independence,

$$(\alpha_i - \beta_i) = 0 \quad \text{for all } i = 0,1,\ldots,m. \qquad \square$$

<u>Lemma (L - 3.2)</u>: Let S_δ and D_δ be operators on $B(R)$ as defined before, and $\{\sigma_m \; ; \; m = 0,1,\ldots,M\}$ be a set of M+1 real constants. Then:

1) There exists an unique set of M+1 real coefficients $\{s_m \; ; \; m = 1,2,\ldots,M+1\}$ such that

$$\sum_{m=0}^{M} \sigma_m \, D_\delta^m = \sum_{m=1}^{M+1} s_m \, S^{m-\tilde{M}-1}.$$

2) Moreover,

$$\sum_{m=1}^{M+1} s_m = \sigma_o$$

3) and each of theese coefficients is given in terms of $\{\sigma_m \; ; \; m = 0,1,\ldots,M\}$

such that

$$D_\delta^m = \sum_{i=0}^{m} \alpha_i \, S_\delta^{i-\tilde{m}} \quad .$$

Now suppose also that

$$D_\delta^m = \sum_{i=0}^{m} \beta_i \, S_\delta^{i-\tilde{m}} \quad .$$

Then

$$0 = D_\delta^m - D_\delta^m = \sum_{i=0}^{m} (\alpha_i - \beta_i) \, S_\delta^{i-\tilde{m}}$$

by independence,

$$(\alpha_i - \beta_i) = 0 \quad \text{for all } i = 0,1,\ldots,m. \qquad \square$$

<u>Lemma (L - 3.2)</u>: Let S_δ and D_δ be operators on $B(R)$ as defined before, and $\{\sigma_m \; ; \; m = 0,1,\ldots,M\}$ be a set of M+1 real constants. Then:

1) There exists an unique set of M+1 real coefficients $\{s_m \; ; \; m = 1,2,\ldots,M+1\}$ such that

$$\sum_{m=0}^{M} \sigma_m \, D_\delta^m = \sum_{m=1}^{M+1} s_m \, S^{m-\tilde{M}-1} .$$

2) Moreover,

$$\sum_{m=1}^{M+1} s_m = \sigma_0$$

3) and each of theese coefficients is given in terms of $\{\sigma_m \; ; \; m = 0,1,\ldots,M\}$

as follows:

$$s_{i+\tilde{M}+1} = \sum_{m=m_s(i)}^{M} (-1)^{\bar{m}-i} \binom{m}{i+\tilde{m}} \frac{\sigma_m}{\delta^m}$$

where

$$m_s(i) = \begin{cases} -2i & ; \text{ if } i \le -1 \\ 0 & ; \text{ if } i = 0 \\ 2i-1 & ; \text{ if } i \ge 1 \end{cases}$$

for each $i = -\tilde{M}, -\tilde{M}+1, \ldots, \bar{M}$.

<u>Proof</u>:

1) By (L - 3.1):

$$\sigma_m D_\delta^m = \frac{\sigma_m}{\delta^m} \sum_{i=0}^{m} (-1)^{m-i} \binom{m}{i} S_\delta^{i-\tilde{m}} =$$

$$= \frac{\sigma_m}{\delta^m} \sum_{i=-\tilde{m}}^{\bar{m}} (-1)^{\bar{m}-i} \binom{m}{i+\tilde{m}} S_\delta^i \quad . \tag{4}$$

Hence

$$\sum_{m=0}^{M} \sigma_m D_\delta^m = \sum_{m=0}^{M} \frac{\sigma_m}{\delta^m} \sum_{i=-\tilde{m}}^{\bar{m}} (-1)^{\bar{m}-i} \binom{m}{i+\tilde{m}} S_\delta^i =$$

$$= \sum_{i=-\tilde{M}}^{\bar{M}} r_i S_\delta^i \tag{5}$$

for some set of M+1 coefficients $\{r_i \; ; \; i = -\tilde{M}, \ldots, \bar{M}\}$. Moreover, this representation of $\sum_{m=0}^{M} \sigma_m D_\delta^m$ in terms of $\{S_\delta^{-\tilde{M}}, \ldots, S_\delta^{\bar{M}}\}$ must be unique by independence of $\{S_\delta^{-\tilde{M}}, \ldots, S_\delta^{\bar{M}}\}$. Then defining

$$s_i = r_{i-\tilde{M}-1}$$

we have proved the first part.

2) Let ${}_{6_m}D^m_\delta$ be written in the following form:

$$ {}_{6_m}D^m_\delta = \sum_{i=0}^{m} \rho_m(i)\, S_\delta^{i-\tilde{m}} \quad . \tag{6}$$

So, by (L - 3.1), we have

$$\rho_m(i) = \frac{6_m}{\delta^m}(-1)^{m-i}\binom{m}{i} \quad .$$

Since

$$\sum_{i=0}^{m}(-1)^{m-i}\binom{m}{i} = (-1)^m \sum_{i=0}^{m}(-1)^i\binom{m}{i} = \begin{cases} 1 ; & \text{if } m = 0 \\ 0 ; & \text{if } m \neq 0 \end{cases}$$

we get

$$\sum_{i=0}^{m} \rho_m(i) = \begin{cases} 6_0 ; & \text{if } m = 0 \\ 0 ; & \text{if } m \neq 0 \end{cases} \quad . \tag{7}$$

Then, by (6) and the result of the first part, we have

$$\sum_{m=0}^{M} 6_m D^m_\delta = \sum_{m=0}^{M}\sum_{i=0}^{m} \rho_m(i)\, S_\delta^{i-\tilde{m}} = \sum_{m=1}^{M+1} s_m S_\delta^{m-\tilde{M}-1} .$$

Now, using (7), it comes:

$$\sum_{m=1}^{M+1} s_m = \sum_{m=0}^{M}\left(\sum_{i=0}^{m} \rho_m(i)\right) = 6_0 \quad .$$

3) Finally let us return to the equation (4) and (5). For each $m = 0,\dots,M$ the coefficient of S^i_δ in (4) is

$$\frac{6_m}{\delta^m}(-1)^{\bar{m}-i}\binom{m}{i-\tilde{m}} \quad ,$$

and for each $i = -\tilde{M},\dots,\bar{M}$ we have:

i) If $i = 0$; there exists r_o in (5) for all $m = 0,\ldots,M$

ii) If $i \geq 1$; there exists r_i in (5) for all m such that $\bar{m} \geq i$

iii) If $i \leq -1$; there exists r_i in (5) for all m such that $-\tilde{m} \leq i$

Then

$$r_i = \sum_{m=m_s(i)}^{M} \frac{\sigma_m}{\delta^m} (-1)^{\bar{m}-i} \binom{m}{i+\tilde{m}}$$

where

$$m_s(i) = \begin{cases} 0 & ; \text{ if } i = 0 \\ \min\{m : \bar{m}=i\} = \min\{2i-1 , 2i\} = 2i-1 & ; \text{ if } i \geq 1 \\ \min\{m : \tilde{m}=-i\} = \min\{-2i , -2i+1\} = -2i & ; \text{ if } i \leq -1 \end{cases}$$

And the third part is proved since

$$s_i = r_{i-\tilde{M}-1} \; . \qquad \square$$

<u>Extension</u>: Let $x = (x_1,\ldots,x_n) \in R^n$ and define the following operator on $B(R^n)$, the linear space of all bounded real-valued functions on R^n:

$$S^p_{\delta_{x_i}} f(x) = f(x_1,\ldots,x_i+p\delta_{x_i},\ldots,x_n)$$

for any $f \in B(R^n)$ and $p \in R$, where $\delta_{x_i} > 0$ is fixed.

So $S^p_{\delta_{x_i}}$, the "partial delta shift operator", is a natural extension of S^p_δ. Following this idea we can define "partial forward and backward difference operators",

$$\Delta_{\delta_{x_i}} = S_{\delta_{x_i}} - I$$

$$\Delta^{-1}_{\delta_{x_i}} = I - S^{-1}_{\delta_{x_i}}$$

and also a "m+1 th-order partial centered slope operator" on $B(R^n)$:

$$D^{m+1}_{\delta_{x_i}} = \frac{1}{\delta_{x_i}} \Delta^{(-1)^m}_{\delta_{x_i}} D^m_{\delta_{x_i}} \; ; \qquad D^o_{\delta_{x_i}} = I \,.$$

In this way, all the results obtained about "ordinary" operators S^p_δ, Δ_δ, Δ^{-1}_δ, and D^m_δ on $B(R)$, have a direct extension to "partial" operators $S^p_{\delta_{x_i}}$, $\Delta_{\delta_{x_i}}$, $\Delta^{-1}_{\delta_{x_i}}$ and $D^m_{\delta_{x_i}}$ on $B(R^n)$.

3.4 - APPROXIMATING DERIVATIVES

The results of lemmas (L - 3.1) and (L - 3.2) are now applied to approximate certain class of linear PDE as a natural extension of classical second-order finite-difference techniques. To begin with, we present a brief review on a particular type of approximation for ordinary derivatives.

Ordinary derivatives: Let $f(x)$ belong to $C^M(x_a,x_b)$, the space of all bounded real-valued functions on the interval (x_a,x_b) which are continuous and have continuous derivatives up to the Mth-order on (x_a,x_b).

The derivative of $f(x)$ on (x_a,x_b) can be defined in terms of operators Δ_δ, Δ_δ^{-1} and D_δ since

$$\frac{d}{dx}f(x) = \lim_{\delta\to 0} \frac{f(x)-f(x-\delta)}{\delta} = \lim_{\delta\to 0} \frac{f(x+\delta)-f(x)}{\delta} =$$

$$= \lim_{\delta\to 0} \frac{1}{\delta}\, \Delta_\delta^{-1} f(x) = \lim_{\delta\to 0} \frac{1}{\delta} \Delta_\delta f(x) = \lim_{\delta\to 0} D_\delta\, f(x) \,.$$

In the same way we can show by induction (e.g., see [21]-[23]) that higher order derivatives may also be written in terms of operators Δ_δ^m, Δ_δ^{-m} or D_δ^m:

$$\frac{d^m}{dx^m}f(x) = \lim_{\delta\to 0}\frac{1}{\delta^m}\, \Delta_\delta^m f(x) = \lim_{\delta\to 0} \frac{1}{\delta^m}\, \Delta_\delta^{-m} f(x) =$$

$$= \lim_{\delta\to 0}\frac{1}{\delta^m}\left[\Delta_\delta^{(-1)^{m-1}}\Delta_\delta^{(-1)^{m-2}}\ldots\Delta_\delta^{-1}\Delta_\delta\right] f(x) =$$

$$= \lim_{\delta\to 0} D_\delta^m f(x) \,.$$

So forward, backward and centered operators (Δ_δ^m, Δ_δ^{-m} and D_δ^m) can be used to approximate derivatives at $x_o \in (x_a,x_b)$, for a sufficiently small δ, as follows:[7]

$$\left.\frac{d^m}{dx^m}f(x)\right|_{x=x_o} \simeq \begin{cases} \frac{1}{\delta^m}\Delta_\delta^{-m} f(x_o) & : \text{(backward operators)} \\ D_\delta^m f(x_o) & : \text{(centered operators)} \\ \frac{1}{\delta^m}\Delta_\delta^m f(x_o) & : \text{(forward operators)} \end{cases} .$$

---

[7] For details concerning with this sort of (first-order) approximation see, for example, [21], [22], [24] and [25].

Partial derivatives: A direct extension of the preceding approximation procedure can be done for partial derivatives as summarized below:

Let $x = (x_1, \ldots, x_n) \in \Omega$, a simply connected open set in R^n and $f(x) \in C^M(\Omega)$, the space of all bounded real-valued functions on Ω which are continuous and have continuous partial derivatives up to the M th-order on Ω.

The operators $\Delta^m_{\delta_{x_i}}$, $\Delta^{-m}_{\delta_{x_i}}$ or $D^m_{\delta_{x_i}}$ can be used to write down partial derivatives of $f(x)$ on Ω,

$$\frac{\partial^m}{\partial x_i^m} f(x) = \lim_{\delta_{x_i} \to 0} \frac{1}{\delta^m_{x_i}} \Delta^m_{\delta_{x_i}} f(x) = \lim_{\delta_{x_i} \to 0} \frac{1}{\delta^m_{x_i}} \Delta^{-m}_{\delta_{x_i}} f(x) = \lim_{\delta_{x_i} \to 0} D^m_{\delta_{x_i}} f(x)$$

and so, assuming a sufficiently small δ_{x_i}, we can approximate partial derivatives (of a single independent variable [8]) at $x_o \in \Omega$ by using partial forward, backward and centered operators ($\Delta^m_{\delta_{x_i}}$, $\Delta^{-m}_{\delta_{x_i}}$ and $D^m_{\delta_{x_i}}$) as follows:

$$\left.\frac{\partial^m}{\partial x_i^m} f(x)\right|_{x=x_o} \simeq \begin{cases} \frac{1}{\delta^m_{x_i}} \Delta^m_{\delta_{x_i}} f(x_o) \; ; & \text{(backward operators)} \\ D^m_{\delta_{x_i}} f(x_o) \; ; & \text{(centered operators)} \\ \frac{1}{\delta^m_{x_i}} \Delta^{-m}_{\delta_{x_i}} f(x_o) \; ; & \text{(forward operators)} \end{cases}$$

8 Here we are excluding the case containing cross-terms, such as $\frac{\partial^2}{\partial x_i \, \partial x_j}$, since they will not be of interest in our further studies.

Remark: Although we are not going to use approximation for cross-terms partial derivatives, it is interesting to notice that:

$$\frac{\partial^m}{\partial x_1^{k_1} \dots \partial x_n^{k_n}} f(x)\Big|_{x=x_o} = \lim_{\substack{\delta_{x_1} \to 0 \\ \vdots \\ \delta_{x_n} \to 0}} \frac{1}{\delta_{x_1}^{k_1} \dots \delta_{x_n}^{k_n}} \prod_{i=1}^{n} \Delta_{\delta_{x_i}}^{k_i} f(x_o)$$

where $m = k_1 + \dots + k_n$.

Finally let us consider the particular problem of approximating the linear partial differential operators

$$L_x^M u(x,t) = \sum_{m=1}^{M} \alpha_m(x) \frac{\partial^m}{\partial x^m} u(x,t)$$

$$L_t^N u(x,t) = \sum_{m=0}^{N} \gamma_m(x) \frac{\partial^m}{\partial t^m} u(x,t)$$

introduced in (2).

Let the S_{δ_x}, S_{δ_t}, $D_{\delta_x}^m$ and $D_{\delta_t}^m$ be operators on $B(R^2)$ as defined before. By lemmas (L - 3.1) and (L - 3.2) with $(x_k,t_n) \in \Omega \subset R^2$, the domain of $u(x,t)$, and using centered operators for approximating partial derivatives we get:

$$L_x^M u(x,t_n)\Big|_{x=x_k} = \sum_{m=1}^{M} \alpha_m(x_k)\, D_{\delta_x}^m\, u(x_k,t_n)$$

$$= \sum_{m=1}^{M} \alpha_m(x_k) \frac{1}{\delta_x^m} \sum_{i=0}^{m} (-1)^{m-i} \binom{m}{i} S_{\delta_x}^{i-\tilde{m}}\, u(x_k,t_n) =$$

$$= \sum_{m=1}^{M+1} a_m(x_k)\, S_{\delta_x}^{m-\tilde{M}-1}\, u(x_k,t_n) =$$

$$= \sum_{m=1}^{M+1} a_m(x_k)\, u(x_k+(m-\tilde{M}-1)\delta_x, t_n) \,.$$

In similar way

$$L_t^N u(x_k,t)\Big|_{t=t_n} = \sum_{m=0}^{N} \gamma_m(x_k)\, D_{\delta_t}^m u(x_t,t_n)$$

$$= \sum_{m=0}^{N} \gamma_m(x_k) \frac{1}{\delta_t^m} \sum_{i=0}^{m} (-1)^{m-i} \binom{m}{i} S_{\delta_t}^{i-\tilde{m}}\, u(x_k,t_n) =$$

$$= \sum_{m=1}^{N+1} c_m(x_k)\, S_{\delta_t}^{m-\tilde{N}-1}\, u(x_k,t_n)$$

$$= \sum_{m=1}^{N+1} c_m(x_k)\, u(x_k,t_n+(m-\tilde{N}-1)\delta_t)$$

where the coefficients $a_m(x_k)$ and $c_m(x_k)$ are such as introduced on lemma (L - 3.2), and so (note that $\alpha_0(x) = 0$):

$$\sum_{m=1}^{M+1} a_m(x_k) = 0$$

$$\sum_{m=1}^{N+1} c_m(x_k) = \gamma_0(x_k)$$

Remarks:

1) When using this kind of centered approximation we must be sure that for a given $(x_k,t_n) \in \Omega$, $(x_k+(m'-\tilde{M}-1)\delta_x, t_n+(m''-\tilde{N}-1)\delta_t) \in \Omega$ (or at least in $\bar{\Omega}$, the closure of Ω) for all $m' = 1,\ldots,M+1$ and $m'' = 1,\ldots,N+1$.

2) Denoting

$$a_k^m = a_m(x_k)$$

$$c_k^m = c_m(x_k)$$

$$u_{k+j}(n+i) = u(x_k + j\,\delta_x\,,\, t_n + i\,\delta_t)$$

we get

$$L_x^M\, u(x,t_n)\Big|_{x=x_k} = \sum_{m=1}^{M+1} a_k^m\, u_{k+m-\tilde{M}-1}(n)$$

$$L_t^N\, u(x_k,t)\Big|_{t=t_n} = \sum_{m=1}^{N+1} c_k^m\, u_k(n+m-\tilde{N}-1)$$

This simplified notation is extensively used in the next chapter.

3) The accuracy of this approximation technique increases for lower order models. The literature dealing with the classical finite-difference techniques [7]-[19] presents discussions on this subject, extressing the accuracy for models with $M = 2$ and $N = 2$.

PART III: STOCHASTIC APPROXIMATION

Stochastic Approximation is a recursive scheme which can be used for parametric estimation in a stochastic environment. Its origins are the works of Robbins and Monro [26], Kiefer and Wolfowitz [27], Blum [28] and the unified general approach given by Dvoretzky [29]. Presently there is a great deal of literature on this subject, both from theoretical and practical viewpoints. Some complete books on stochastic approximation have already been published [30], [31], and interesting surveys regarding mainly the applications are also available [32]-[42], [75].

This technique has been extensively used for identification purposes in memoryless systems and LPS (e.g., see [43]-[54]) and also, but not so extensively, for DPS identification [55]-[60]. Concerning the literature in system identification by stochastic approximation, it has become common practice to refer to Dvoretzky's work [29] and then to proceed directly to applicable algorithms. But it happens that this "bridge" linking theory and practice is not so obvious and the gap between them is not so narrow either.

For this reason we begin this third part on mathematical preliminaries by presenting the Dvoretzky theorem in abstract spaces. In an attempt to bridge that gap between theory and practice, we introduce some applicable stochastic approximation algorithms in Hilbert spaces. Particular cases involving algorithms for operators on finite-dimensional spaces are also included, since they will be required in the next chapter.

3.5 - THE DVORETZKY'S THEOREM IN BANACH SPACES

Wolfowitz [61]and Derman and Sacks [62] presented new proofs for the Dvoretzky theorem. In both cases, the proofs are given for real-valued random variables and in [62] they also presented an extension for finite-dimensional random vectors. Schmetterer [63] considered stochastic approximation algorithms, in particular the unified Dvoretzky's approach, in Hilbert spaces. The previous works were generalized by Venter [64], who proposed a wider class of algorithms in Hilbert spaces.

In his original paper, Dvoretzky [29] formulated an infinite-dimensional version for stochastic approximation algorithms in normed linear spaces, whose proof is a natural extension of the scalar (real) case. This theorem is formulated below in a simplified version, and the reader interested in its proof is referred to [29].

Theorem (T - 3.1) (Dvoretzky): Assume $x_0 (\|x_0\| < \infty)$ a fixed point in B, a Banach space. Let $\{z(n) \; ; \; n = 0,1,2,\ldots\}$ be a B-valued random sequence, such that

$$\sum_{n=0}^{\infty} E\{\|z(n)\|^2\} < \infty \tag{8}$$

where $\| \; \|$ stands for the norm in B. Let x(0) be a B-valued second-order random variable, and consider the following algorithm (a discrete-time dynamical system) in B:

$$x(n+1) = T_n x(n) + z(n) \tag{9}$$

where $\{T_n : B \to B \; ; \; n = 0,1,2,\ldots\}$ is a family of bounded operators on B.

If:

$$\|T_n x - x_o\| \leq F_n \|x - x_o\| \tag{10}$$

for any $x \in B$, and

$$E\{\|T_n x(n) - x_o + z(n)\|^2\} \leq E\{\|T_n x(n) - x_o\|^2\} + E\{\|z(n)\|^2\} \tag{11}$$

for all $n = 0,1,2,\ldots$, where $\{F_n ; \ n = 0,1,2,\ldots\}$ is a real sequence such that:

$$F_n > 0 \tag{12}$$

for all n, and

$$\prod_{n=0}^{\infty} F_n = 0 \quad . \tag{13}$$

Then $\{x(n)\}$ converges to x_o in quadratic mean (q.m.) and with probability one (w.p.1): [9]

$$\lim_{n \to \infty} E\{\|x(n) - x_o\|^2\} = 0$$

and

$$P\{\lim_{n \to \infty} x(n) = x_o\} = 1 \quad .$$

---

9 The concepts of convergence "with probability one" (w.p.1), "almost certainly" (a.c.), "almost sure (or surely)" (a.s.), and "almost everywhere" (a.e.) are equivalent. For theoretical consideration see, for example, [1], [65]-[69].

Special case (S - 3.1): Let B = H, a Hilbert space, and let $\langle \,;\, \rangle$ denote the inner product in H. Anyone of the (sufficient) conditions stated below can be used to substitute the condition (11) in the thoorem:

$$E\left\{z(n) \,\middle|\, x(0), z(0), \ldots, z(n-1)\right\} = 0 \qquad \text{w.p.1} \qquad (11')$$

for all n, or

$$\sum_{n=0}^{\infty} E\left\{\left|\langle T_n x(n) - x_o \,;\, z(n)\rangle\right|\right\} < c < \infty . \qquad (11'')$$

Proof:

$$E\left\{\|T_n x(n) - x_o + z(n)\|^2\right\} =$$

$$= E\left\{\|T_n x(n) - x_o\|^2\right\} + E\left\{\|z(n)\|^2\right\} +$$

$$+ E\left\{\langle T_n x(n) - x_o \,;\, z(n)\rangle\right\} + E\left\{\langle z(n) \,;\, T_n x(n) - x_o\rangle\right\} \leq$$

$$\leq E\left\{\|T_n x(n) - x_o\|^2\right\} + E\left\{\|z(n)\|^2\right\} + 2\,E\left\{\left|\langle T_n x(n) - x_o \,;\, z(n)\rangle\right|\right\} .$$

(11'): $$E\left\{\langle T_n x(n) - x_o \,;\, z(n)\rangle\right\} = E\left\{E\left\{\langle T_n x(n) - x_o \,;\, z(n)\rangle \,\middle|\, x(n)\right\}\right\} =$$

$$= E\left\{\left\langle T_n x(n) - x_o \,;\, E\left\{z(n) \,\middle|\, x(n)\right\}\right\rangle\right\} =$$

$$= E\left\{\left\langle T_n x(n) - x_o \,;\, E\left\{z(n) \,\middle|\, x(0), z(0), \ldots, z(n-1)\right\}\right\rangle\right\} = 0 .$$

In the same way:

$$E\left\{\langle z(n) \,;\, T_n x(n) - x_o\rangle\right\} = 0 .$$

Special case (S - 3.1): Let $B = H$, a Hilbert space, and let $\langle\ ;\ \rangle$ denote the inner product in H. Anyone of the (sufficient) conditions stated below can be used to substitute the condition (11) in the theorem:

$$E\{z(n)\,|\,x(0),z(0),\dots,z(n-1)\} = 0 \qquad \text{w.p.1} \qquad (11')$$

for all n, or

$$\sum_{n=0}^{\infty} E\{|\langle T_n x(n) - x_o\ ;\ z(n)\rangle|\} < c < \infty . \qquad (11'')$$

Proof:

$$E\{\|T_n x(n) - x_o + z(n)\|^2\} =$$

$$= E\{\|T_n x(n) - x_o\|^2\} + E\{\|z(n)\|^2\} +$$

$$+ E\{\langle T_n x(n) - x_o\ ;\ z(n)\rangle\} + E\{\langle z(n)\ ;\ T_n x(n) - x_o\rangle\} \leq$$

$$\leq E\{\|T_n x(n) - x_o\|^2\} + E\{\|z(n)\|^2\} + 2\,E\{|\langle T_n x(n) - x_o\ ;\ z(n)\rangle|\} .$$

(11'): $$E\{\langle T_n x(n) - x_o\ ;\ z(n)\rangle\} = E\{E\{\langle T_n x(n) - x_o\ ;\ z(n)\rangle\,|\,x(n)\}\} =$$

$$= E\{\langle T_n x(n) - x_o\ ;\ E\{z(n)\,|\,x(n)\}\rangle\} =$$

$$= E\{\langle T_n x(n) - x_o\ ;\ E\{z(n)\,|\,x(0),z(0),\dots,z(n-1)\}\rangle\} = 0 .$$

In the same way:

$$E\{\langle z(n)\ ;\ T_n x(n) - x_o\rangle\} = 0 .$$

(11"): Set

$$\sigma_n^2 = E\{\|z(n)\|^2\} + 2\,E\{|\langle T_n x(n) - x_o \; ; \; z(n)\rangle|\} \; .^{10}$$

So

$$E\{\|T_n x(n) - x_o + z(n)\|^2\} \neq E\{\|T_n x(n) - x_o\|^2\} + \sigma_n^2$$

where

$$\sum_{n=0}^{\infty} \sigma_n^2 < \infty$$

and that is sufficient to insure the theorem (for discussions, see [29] and [64]). □

Special case (S - 3.2): Now set $B = BL(R^k)$, the class of all bounded linear operators from R^k into itself. Let A^* and $tr[A]$ denote the adjoint and the trace of an element A in $BL(R^k)$, respectively. As in (S - 3.1), the conditions below can be used to substitute (11) in (T - 3.1):

$$E\{Z(n) \mid X(0), Z(0), \ldots, Z(n-1)\} = 0 \qquad \text{w.p.1}$$

for all n, or

$$\sum_{m=0}^{\infty} E\{tr[(T_n X(n) - X_o)^* Z(n)]\} < c < \infty .$$

---

10 If H is a real Hilbert space, the absolute valued appearing in (11") may be omited.

Proof:

$BL(R^k)$, with addition and scalar multiplication defined as usual, is a Hilbert space with the inner product defined as

$$\langle A ; B \rangle = tr[A^* B]$$

for all A, B in $BL(R^k)$ [70]. So, by invariance of the trace (which is real) [71] and by equivalence of norms [72] in $BL(R^k)$, the proof follows as in (S - 3.1). $\square$

3.6 - A PROPOSED STOCHASTIC APPROXIMATION ALGORITHM IN HILBERT SPACE

In this section we present two corollaries for the preceding theorem. A similar version (in R^1) can be found in [37], where the proof stating the connection with the Dvoretzky theorem is somewhat obscure.

First we need to prove the following lemma:

Lemma (L - 3.3): Let $\{\zeta_n \; ; \; n = 0,1,2,\dots\}$ be a real positive sequence such that

$$\frac{\zeta_n}{\zeta_{n+1}} \leq 1 + \zeta_n + \xi_n$$

for all n, where:

$$\xi_n \geq 0 \quad ; \quad \sum_{n=0}^{\infty} \xi_n < \infty .$$

Define

$$\Phi(n,i) = \prod_{j=i}^{n-1} (1-\zeta_j) \; ; \qquad i < n$$

$$\Phi(n,n) = 1$$

Then:

$$\zeta_i \, \Phi(n+1,i+1) \leq k \, \zeta_n$$

for some finite positive constant k.

<u>Proof</u>:

$$\zeta_i \, \Phi(n+1,i+1) = \zeta_i \prod_{j=i+1}^{n} (1-\zeta_j) = \zeta_i \prod_{j=i+1}^{n} \zeta_j \left(\frac{1}{\zeta_j} - 1\right) =$$

$$= \prod_{j=i}^{n-1} \zeta_j \left(\frac{1}{\zeta_{j+1}} - 1\right) \zeta_n = \zeta_n \prod_{j=i}^{n-1} \left(\frac{\zeta_j}{\zeta_{j+1}} - \zeta_j\right) \leq$$

$$\leq \zeta_n \prod_{j=i}^{n-1} (1+\xi_j) \leq \zeta_n \lim_{n\to\infty} \prod_{j=i}^{n-1} (1+\xi_j) =$$

$$= \zeta_n k \; ; \qquad 0 < k < \infty$$

since [73]:

$$\xi_n \geq 0 \quad \text{and} \quad \sum_{n=0}^{\infty} \xi_n < \infty \longleftrightarrow 0 < \prod_{n=n_o}^{\infty} (1+\xi_n) = k < \infty \, . \qquad \Box$$

<u>Corollary (C - 3.1)</u>: Let B = H be a Hilbert space, x_o a fixed point in H ($\|x_o\| < \infty$), and consider the following algorithm in H:

$$x(n+1) = (1-\zeta_n)\, x(n) + \zeta_n \, y(n) \tag{14}$$

where $\{\zeta_n \; ; \; n = 0,1,2,\ldots\}$ is a real sequence, $\{y(n) \; ; \; n = 0,1,2,\ldots\}$ ia a H-valued random sequence, and $x(0)$ is a second-order H-valued random variable independent of $\{y(n)\}$.

If:

i) $\zeta_n \in (0,1)$

ii) $\sum_{n=0}^{\infty} \zeta_n = \infty$

iii) $\sum_{m=0}^{\infty} \zeta_n^2 < \infty$

iv) $E\{\|y(n)\|^2\} < \sigma^2 < \infty$

and

v-a) $\{\zeta_n\}$ as in (L - 3.3) [11]

v-b) $E\{y(n)\} = x_o$; for all n

v-c) $\sum_{n=0}^{m-1} E\{|\langle y(i) - x_o \; ; \; y(n) - x_o\rangle|\} < c < \infty$

or

v') $E\{y(\bar{n}) - x_o \,|\, y(0),\ldots,y(n-1)\} = 0$ w.p.1

for all n. [12]

---

11 For example, the sequence

$$\zeta_n = \frac{1}{(an+b)^{\alpha}} \; ; \qquad 0 < a \leq 1 \; ; \quad b > 1 \; ; \quad \tfrac{1}{2} < \alpha \leq 1$$

satisfies (i) - (iii), and (v-a) [73].

12 Note that, from (v') we get (v-b) [65], [66].

Then:

$$\lim_{n\to\infty} E\{\|x(n) - x_o\|^2\} = 0$$

and

$$P\{\lim_{n\to\infty} x(n) = x_o\} = 1 \; .$$

Proof:

By (14)

$$x(n+1) = (1-\zeta_n)\left[x(n) - x_o\right] + x_o + \zeta_n\left[y(n) - x_o\right] .$$

Set

$$F_n = (1-\zeta_n)$$

$$z(n) = \zeta_n\left[y(n) - x_o\right]$$

and define operators in H, as follows:

$$T_n x = F_n(x - x_o) + x_o$$

for any x in H. So we get an algorithm as in (9):

(9): $\qquad x(n+1) = T_n x(n) + z(n)$

where the condition in (8), (10), (12) and (13) are satisfied:

(8): $$\sum_{n=0}^{\infty} E\{\|z(n)\|^2\} = \sum_{n=0}^{\infty}\zeta_n^2\left(E\{\|y(n)\|^2\} - \|x_o\|^2\right) <$$

$$< \left(\sigma^2 - \|x_o\|^2\right)\sum_{n=0}^{\infty}\zeta_n^2 < \infty .$$

(10): $\qquad \|T_n x - x_o\| = F_n\|x - x_o\|$, $\qquad$ for any x in H.

(12): $\qquad 0 < F_n < 1$, $\qquad$ since $0 < \zeta_n < 1$.

(13): $\prod_{n=0}^{\infty} F_n = 0$, since[13] $0 < \zeta_n < 1$ and $\sum_{n=0}^{\infty} \zeta_n = \infty$.

Then, in order to complete the proof, we just need to verify the condition (11) - theorem (T - 3.1) - or one of its equivalent forms, (11') or (11"), as in (S - 3.1). In the following we show that (v') $\rightarrow$ (11') and (v) $\rightarrow$ (11")

(11'):
$$E\{z(n) \mid x(0), z(0), \ldots, z(n-1)\} =$$
$$= \zeta_n E\{y(n) - x_o \mid x(0), y(0), \ldots, y(n-1)\} =$$
$$= \zeta_n E\{y(n) - x_o \mid y(0), \ldots, y(n-1)\} =$$
$$= 0 \qquad \text{w.p.1} , \quad \forall n$$

since $x(0)$ is independent of $\{y(n)\}$.

(11"): Let us rewrite the algorithm in (14):

$$x(n+1) - x_o = (1 - \zeta_n)[x(n) - x_o] + \zeta_n [y(n) - x_o] .$$

So, for any fixed m, we have:

$$c_{xy}(n+1,m) \leq \Phi(n+1,n)\ c_{xy}(n,m) + \zeta_n\ c_{yy}(n,m) \tag{15}$$

where

$$c_{xy}(n,m) = E\{|\langle x(n) - x_o \ ; \ y(m) - x_o \rangle|\}$$

$$c_{yy}(n,m) = E\{|\langle y(n) - x_o \ ; \ y(m) - x_o \rangle|\}$$

$$\Phi(n,i) = \prod_{j=i}^{n-1} (1 - \zeta_j) \ ; \quad i < n$$

$$\Phi(n,n) = 1$$

---

[13] See, for example, [74] pp. 146.

On iterating the inequality (15) we get:

$$c_{xy}(n,m) \leq \Phi(n,0)\ c_{xy}(0,m) + \sum_{i=0}^{n-1} \Phi(n,i+1)\ \zeta_i\ c_{yy}(i,m) .$$

Since $x(0)$ is independent of $\{y(n)\}$ and $E\{y(n) - x_o\} = 0$ for all n,

$$c_{xy}(0,m) = 0 , \qquad \text{for any } m.$$

Thus, setting $m = n$, we have:

$$c_{xy}(n,n) \leq \sum_{i=0}^{n-1} \Phi(n,i+1)\ \zeta_i\ c_{yy}(i,n) .$$

Now note that:

$$\sum_{i=0}^{n-1} c_{yy}(i,n) < c < \infty \qquad \text{(by assumption (v-c))}$$

and

$$\Phi(n+1,i+1)\ \zeta_i \leq k\ \zeta_n$$

for some finite positive constant k (by assumption (v-a) and lemma (L - 3.3)).

Then:

$$\sum_{n=0}^{\infty} E\{|\langle T_n x(n) - x_o\ ;\ z(n)\rangle|\} = \sum_{n=0}^{\infty} F_n\ \zeta_n\ c_{xy}(n,n) \leq$$

$$\leq \sum_{n=0}^{\infty} \zeta_n (1 - \zeta_n) \sum_{i=0}^{n-1} \Phi(n,i+1)\ \zeta_i\ c_{yy}(i,n) =$$

$$= \sum_{n=0}^{\infty} \zeta_n \sum_{i=0}^{n-1} \Phi(n+1,i+1)\ \zeta_i\ c_{yy}(i,n) \leq$$

$$\leq \sum_{n=0}^{\infty} k\,\zeta_n^2 \sum_{i=0}^{n-1} c_{yy}(i,n) \leq c\,k \sum_{n=0}^{\infty} \zeta_n^2 < \infty$$

and (11") is satisfied. □

Finally we present a particular case of stochastic approximation algorithms for bounded linear operators on R^k.

Corollary (C - 3.2): The corollary (C - 3.1) remains valid if $B = BL(R^k)$, the class of all bounded linear operators from R^k into itself, provided that the condition (v-c) is replaced by

$$\sum_{i=0}^{n-1} E\{tr[(Y(i)-X_o)^*(Y(n)-X_o)]\} < c < \infty .$$

Proof:

Define

$$c_{XY}(n,m) = E\{tr[(X(n)-X_o)^*(Y(m)-X_o)]\}$$
$$c_{XY}(n,m) = E\{tr[(Y(n)-X_o)^*(Y(m)-X_o)]\}$$

and the proof follows exactly as in (C - 3.1) [14], by using the results obtained in (S - 3.2). □

Remarks:

1) In [37] Fu proposed a similar version for stochastic approximation algorithms in R^1, where direct extension to R^k and $BL(R^k)$ can be

[14] In this case the equality in (15) holds.

obtained.[15] In his formulation he did not require the conditions (v-a) and (v-c) or (v'). So, if one belives in [37], a simplified version of (C - 3.1) in R^k and (C - 3.2) can be formulated by assuming only the conditions (i) - (iv) and (v-b). Note that, in this case,

$$z(n) = \zeta_n [y(n) - x_o]$$

$$E\{y(n) - x_o\} = 0 \; ; \quad \forall n$$

$$E\{\|y(n)\|^2\} < \delta^2 < \infty$$

are being used to substitute the original sufficient condition

$$E\{z(n) \,|\, x(0), x(1), \ldots, x(n)\} = 0 \qquad \text{w.p.1} \; ; \quad \forall n \qquad (11')$$

assumed by Dvoretzky in [29] for the scalar (real) case.

2) Finally we remark, as proved by Venter [64], that conditions weaker than (11'), such as

$$\sum_{n=0}^{\infty} \left(E\{\|E\{z(n) \,|\, x(0), z(1), \ldots, z(n-1)\}\|^2\}\right)^{\frac{1}{2}} < \infty \qquad (16)$$

$$\sum_{n=0}^{\infty} \|E\{z(n) \,|\, x(0), z(1), \ldots, z(n-1)\}\| < \infty \qquad \text{w.p.1} \qquad (17)$$

are able to ensure the convergence for algorithms in Hilbert spaces, in quadratic mean and with probability one in case of (16), and with probability one in case of (17). (Note that (11') implies (16) which implies (17). For detailed discussion see [64]).

---

[15] Such extensions to finite-dimensional spaces have already been successfully used for identification purposes in memoryless systems [45] and LPS [47], [49].

CHAPTER 4

IDENTIFICATION FOR A CLASS OF LINEAR DPS USING STOCHASTIC APPROXIMATION

This is the central chapter of this work. It presents a new method for identifying distributed systems operating in a stochastic environment, where no restriction concerning probability distributions is imposed.

A class of linear models driven by random inputs and observed through noisy measurements is considered. These measurements are taken at a limited number of discrete points located in the spatial domain.

The method is classified as class Γ_2: First of all a time-space discretization is applied in order to approximate the infinite-dimensional model (described in terms of a linear PDE) by a finite-dimensional one (described in terms of a linear vector difference-equation). So, higher order finite-difference techniques are used to reduce the DPS to a discrete-time LPS (stage I, path 2-A). Thanks to the linearity in the parameters, the space-varying coefficients are placed in an explicit form, and are then identified by using recursive assymptotically unbiased stochastic approximation algorithms (stage II). The method is suitable for on-line applications and extraneous terms may be included in the original model.

4.1 - PROBLEM FORMULATION

Model description: Consider a DPS modeled by PDE and let $u(x,t) \in H(\Omega)$ denote the dependent variable, as follows:

$$L_t^N u(x,t) = L_x^M u(x,t) + \beta(x)\, w_x(t)$$
$$x \in X = (0,\ell)\ ; \qquad t > 0 \tag{1-a}$$

where the linear partial differential operators

$$L_t^N u(x,t) = \sum_{m=0}^{N} \gamma_m(x) \frac{\partial^m}{\partial t^m} u(x,t)$$
$$L_x^M u(x,t) = \sum_{m=1}^{M} \alpha_m(x) \frac{\partial^m}{\partial x^m} u(x,t) \tag{1-b}$$

are such as introduced in the first part of the last chapter, as well as the appropriate function space $H(\Omega)$, $\Omega = (0,\ell) \times (0,\infty) \subset R^2$.

The input disturbances $\{w_x(t)\ ;\ t \geq 0\}$ are taken to be real-valued second-order stochastic processes for each x in $(0,\ell)$. [1] The real-valued space-varying parameters $\{\alpha_1(x),\ldots,\alpha_M(x)\}$, $\{\gamma_0(x),\ldots,\gamma_N(x)\}$ and $\beta(x)$ are supposed to be in $BV[0,\ell]$, the space of all functions of bounded

---

1 $w_x(t)$ can be thought as an infinite-dimensional vector-valued second-order process. Actually, using a more sophisticated mathematical terminology, $\{w_\xi\ ;\ \xi \in (0,\ell) \times [0,\infty) \subset R^2\}$ is a random field, rather than a stochastic process [1]. The existence and uniqueness of solutions for such stochastic DPS has been fully investigated by Curtain and Falb [2], [3].

variation[2] on the interval $[0,\ell]$; and $|\gamma_N(x)| > \gamma > 0$, $\beta(x) \neq 0$, for all x in $(0,\ell)$.[3]

Observations: Assume that noisy observations are available over an equidistant[4] partition P_x of $\bar{X} = [0,\ell]$.

$$z(x_k,t) = h\,u(x_k,t) + d\,v(t) \; ; \qquad t > 0 \tag{2}$$

where the observation noise $\{v(t) \; ; \; t > 0\}$ is taken to be a real-valued second-order stochastic process, and h, d are real constants $(h \neq 0)$.

Problem statement: To identify the set of M parameter functions $\{\alpha_m : [0,\ell] \to R \; ; \; m = 1,\ldots,M\}$ appearing in the spatial-differential operator L_x^M, based on noisy observations $z(x_k,t)$. Under this formulation, the solution lies in $BV[0,\ell]$, the infinite-dimensional parameter space. We consider here a finite-dimensional version:

---

2 By a partition $P_{[a,b]}$ of the interval $[a,b]$, we mean a finite set of points $x_k \in [a,b]$, $k = 0,1,\ldots,K$, such that $a = x_0 < x_1 < \ldots < x_K = b$. A function f defined on $[a,b]$ is said to be of bounded variation if there is a constant f_o so that for any partition of $[a,b]$

$$\sum_{k=0}^{K-1} |f(x_{k+1}) - f(x_k)| < f_o < \infty$$

3 Since L_x^M and L_t^N are defined for $x \in (0,\ell)$, the values of α_m, γ_m and β in $x=0$ and $x=\ell$ have no significance for us. To ensure that these functions are in $BV[0,\ell]$ we can define any real value for them in $x=0$ and $x=\ell$.

4 A partition is said to be "equidistant", if $x_{k+1} - x_k$ is constant for all $k = 0,1,\ldots,K-1$.

"Identify $\{\alpha_m \; ; \; m=1,\ldots,M\}$ for points $x_k \in P_x$". Since $\alpha_m \in BV[0,\ell]$, it is always possible to find an equidistant partition P_x of $[0,\ell]$ such that the finite sequence $\{\alpha_m(x_k) \; ; \; x_k \in P_x\}$ is a good (the goodness depending on how small we choose $\delta_x = x_{k+1} - x_k$) approximation of $\alpha_m(x)$ $x \in [0,\ell]$, for all $m=1,\ldots,M$.

<u>Boundary and initial conditions</u>: A complete description of a physical DPS requires more information than is provided by the distributed model in (1). It is necessary to add some supplementary relations: initial (and/or terminal) and boundary conditions.[5]

Let a set of initial conditions (IC) for the DPS modeled by (1) be given by:

$$\frac{\partial^i}{\partial t^i} u(x,t)\Big|_{t=0} = g_i(x) \; ; \quad x \in [0,\ell]; \quad i=0,1,\ldots,N-1 \qquad (3)$$

where the real-valued initial functions $g_i(x)$ are bounded and continuous on $[0,\ell]$.

Although the initial conditions written in (3) are not unique, they are quite representative for the great majority of physical systems modeled by (1). The same does not happen with a set of boundary conditions. Different experiments on the same system (e.g., one modeled by (1)), provide us with different types of "a priori" information which are expressed by different sets of boundary conditions.

---

[5] We reserve the term "boundary condition" for conditions given only at the spatial boundary (in our case, at $x=0$ and $x=\ell$).

"Identify $\{\alpha_m \; ; \; m=1,\ldots,M\}$ for points $x_k \in P_x$". Since $\alpha_m \in BV[0,\ell]$, it is always possible to find an equidistant partition P_x of $[0,\ell]$ such that the finite sequence $\{\alpha_m(x_k) \; ; \; x_k \in P_x\}$ is a good (the goodness depending on how small we choose $\delta_x = x_{k+1} - x_k$) approximation of $\alpha_m(x)$ $x \in [0,\ell]$, for all $m=1,\ldots,M$.

Boundary and initial conditions: A complete description of a physical DPS requires more information than is provided by the distributed model in (1). It is necessary to add some supplementary relations: initial (and/or terminal) and boundary conditions.[5]

Let a set of initial conditions (IC) for the DPS modeled by (1) be given by:

$$\frac{\partial^i}{\partial t^i} u(x,t)\Big|_{t=0} = g_i(x) \; ; \quad x \in [0,\ell]; \quad i=0,1,\ldots,N-1 \qquad (3)$$

where the real-valued initial functions $g_i(x)$ are bounded and continuous on $[0,\ell]$.

Although the initial conditions written in (3) are not unique, they are quite representative for the great majority of physical systems modeled by (1). The same does not happen with a set of boundary conditions. Different experiments on the same system (e.g., one modeled by (1)), provide us with different types of "a priori" information which are expressed by different sets of boundary conditions.

---

[5] We reserve the term "boundary condition" for conditions given only at the spatial boundary (in our case, at $x=0$ and $x=\ell$).

Since our goal is the identification of parameters appearing only in the distributed model described by the dynamic equation (1), we have two possible ways to proceed: 1) We can assume a particular set of boundary conditions which is representative for many physical experiments in DPS. This approach can also be thought as if we could choose the experimental conditions (concerning with the boundary) on which the DPS would work for identification purpose. 2) The second way would be to assume a general (but not explicit) partial differential equation operating at the boundary.

We choose the first way because it will permit us to carry out a general identification procedure, without especifying the order of the operators L_x^M and L_t^N, up to the point of computational implementation.

For simplicity assume a set of homogeneous boundary conditions (BC):

$$\begin{aligned} &\frac{\partial^i}{\partial x^i} u(x,t)\Big|_{x=0} = 0 \; ; \qquad t \geq 0 \; ; \qquad i = 0,1,\ldots,\bar{M}-1 \\ &\frac{\partial^i}{\partial x^i} u(x,t)\Big|_{x=\ell} = 0 \; ; \qquad t \geq 0 \; ; \qquad i = 0,1,\ldots,\bar{M}-1 \end{aligned} \tag{4}$$

Remarks:

1) (Nonhomogeneous boundary conditions) The identification method that will be proposed here does not require homogeneous BC as a necessary condition for its applicability. We could have assumed nonhomogeneous boundary conditions such as

$$\left.\frac{\partial^i}{\partial x^i} u(x,t)\right|_{x=0} = h_i^o(t) \ ; \qquad t \geq 0 \ ; \qquad i = 0,1,\ldots,\tilde{M}-1$$

$$\left.\frac{\partial^i}{\partial x^i} u(x,t)\right|_{x=\ell} = h_i^\ell(t) \ ; \qquad t \geq 0 \ ; \qquad i = 0,1,\ldots,\bar{M}-1 \tag{4'}$$

where $h_i^o(t)$ and $h_i^\ell(t)$ may be deterministic or random real-valued functions. Remarks will be made along this chapter in order to show that nonhomogeneous boundary conditions may also be considered.

2) (Random initial conditions) We could have assumed the initial functions $g_i(x)$ as random functions rather than deterministic ones. But this assumption would not bring any further generalization to our identification method, wich already assumes random inputs. Also note that, in case of homogeneous boundary conditions, we must have $g_o(0) = g(\ell) = 0$.

4.2 - REDUCTION TO A FINITE-DIMENSIONAL STATE SPACE

Space-time discretization: The space and time-domain can be partitioned as follows:

1) Discretization of space-domain $\bar{X} = \{x : 0 \leq x \leq \ell\}$:

Define

$$k \in \bar{S} = \{0,1,\ldots,K+M-1\}$$

where:

i) The integer M is the order of the operator L_x^M.

ii) $K = \dfrac{\ell}{\delta_x} - M + 1$.

iii) The real constant $\delta_x > 0$ is such that K is an integer $\geq$ M+1 .

So, the function

$$x_k = k\,\delta_x$$

from $\bar{S}$ into $\bar{X}$ defines a partition P_x of $\bar{X}$. The set $\bar{S}$ is called "the discrete space-domain".

2) Discretization of time-domain $\{t : t \geq 0\}$:

Define

$$n \in \bar{T} = \{0,1,\ldots\} .$$

A function from $\bar{T}$ into $[0,\infty)$ such as

$$t_n = n\,\delta_t$$

where $\delta_t > 0$ is a real constant, defines a partition P_t of the interval $[0,\infty)$. The set $\bar{T}$ is called "the discrete time-domain".

<u>Remarks</u>:

1) The sets

$$S = \{\tilde{M},\tilde{M}+1,\ldots,K+\tilde{M}-1\} \subset \bar{S}$$

$$\mathring{S} = \{2\tilde{M},2\tilde{M}+1,\ldots,K+\tilde{M}-\bar{M}-1\} \subset S$$

$$T' = \{\tilde{N},\tilde{N}+1,\ldots\} \subset \bar{T}$$

$$T = \{N,N+1,\ldots\} \subset T'$$

$$\mathring{T} = \{N+\tilde{N},N+\tilde{N}+1,\ldots\} \subset T$$

will be of particular interest in our further studies. The discretization procedure, as well as the location of S, $\mathring{S}$, T', T and $\mathring{T}$, are shown in figures 4 and 5.

2) The condition $K \geq M+1$ imposed to δ_x, ensures that S and $\mathring{S}$ have at least $M+1$ and 1 points, respectively.

3) For reasons that will become clear later in this section, we assume δ_t such that

$$- \delta_t \, \gamma_{N-1}(x) \neq \gamma_N(x) \; ; \qquad \forall \, x \in X$$

if $N \in Z_e$. This condition is always attainable since $|\gamma_N(x)| > \gamma > 0$ and $\gamma_{N-1}(x)$ is bounded $\forall \, x \in X$.

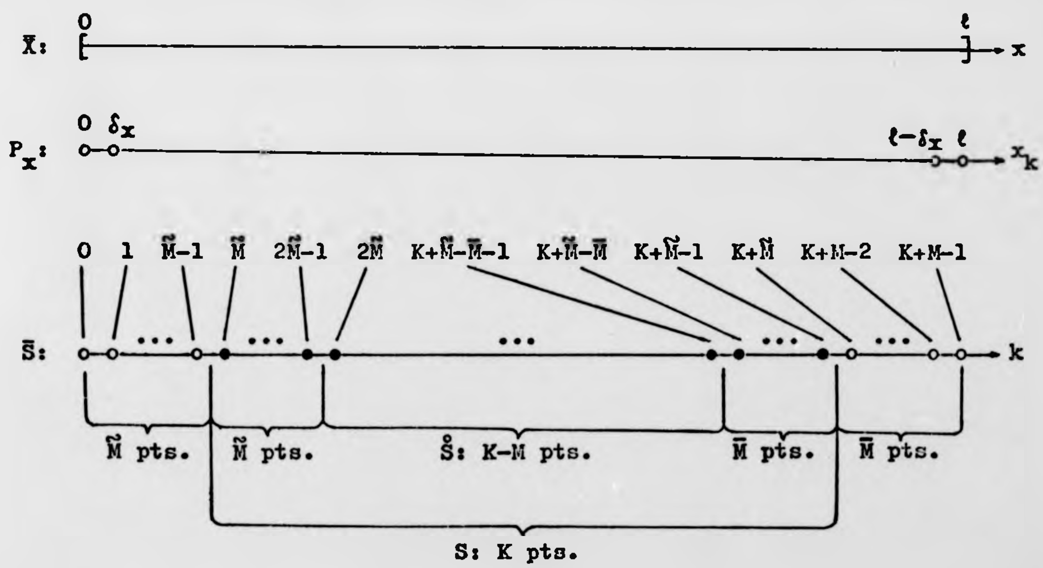

Fig. 4: Discretization of Space-Domain.

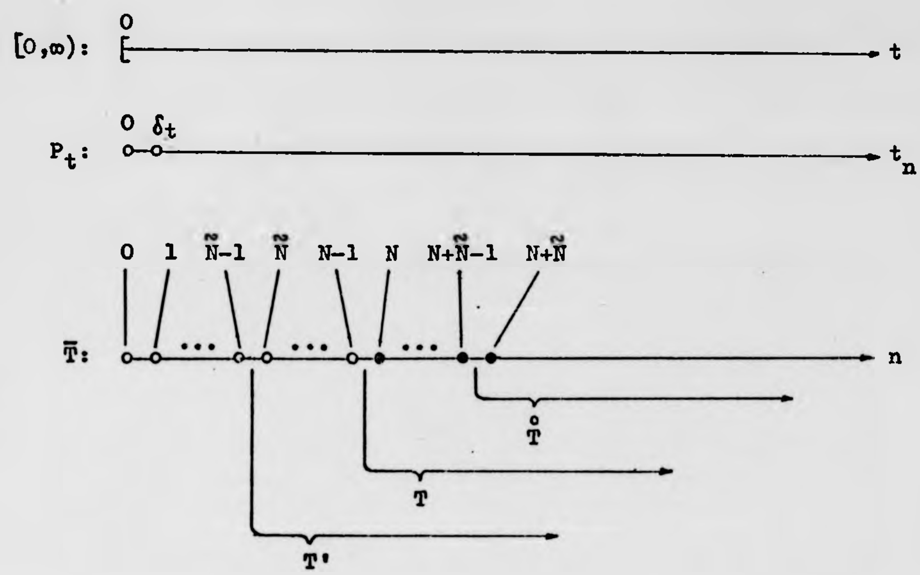

Fig. 5: Discretization of Time-Domain.

<u>Approximating partial derivatives</u>: Now we use space and time shift operators

$$S_{\delta_x}^{\pm 1}\, u(x_k,t_n) = u(x_k \pm \delta_x, t_n)$$

$$S_{\delta_t}^{\pm 1}\, u(x_k,t_n) = u(x_k, t_n \pm \delta_t) \qquad (x_k,t_n) \quad P_x \times P_t$$

to write down approximations for partial derivatives as introduced in the second part of the last chapter. In this way, the identification problem formulated in an infinite-dimensional state space as in (1)-(4), can be reduced to the following finite-dimensional discrete version (The approximation procedure is illustred on figure 6):

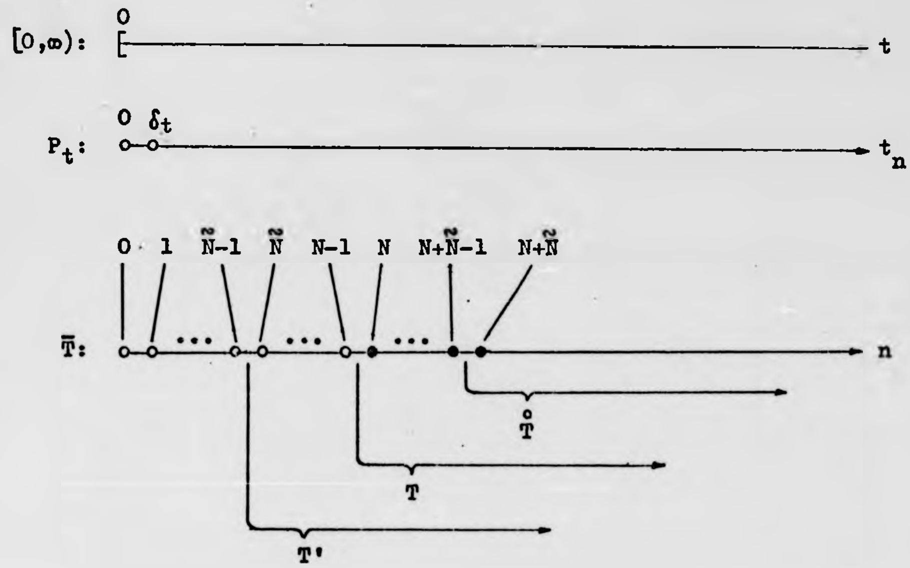

Fig. 5: Discretization of Time-Domain.

Approximating partial derivatives: Now we use space and time shift operators

$$S_{\delta_x}^{\pm 1}\, u(x_k,t_n) = u(x_k \pm \delta_x, t_n)$$
$$(x_k,t_n) \quad P_x \times P_t$$
$$S_{\delta_t}^{\pm 1}\, u(x_k,t_n) = u(x_k, t_n \pm \delta_t)$$

to write down approximations for partial derivatives as introduced in the second part of the last chapter. In this way, the identification problem formulated in an infinite-dimensional state space as in (1)-(4), can be reduced to the following finite-dimensional discrete version (The approximation procedure is illustred on figure 6):

$$
\left\{
\begin{array}{lll}
\text{MODEL:} & \displaystyle\sum_{m=1}^{N+1} c_k^m\, u_k(n+m-\tilde{N}-1) = \sum_{m=1}^{M+1} a_k^m\, u_{k+m-\tilde{M}-1}(n) + \beta_k\, w_k(n) & (5) \\
 & (k,n) \in S \times T' & \\
 & & \\
\text{OBS:} & z_k(n) = h\, u_k(n) + d\, v(n) \;; \qquad (k,n) \in S \times T & (6) \\
 & & \\
\text{IC:} & u_k(0), \ldots, u_k(N-1) \text{ given by } g_i(x_k) \;; \qquad k \in S & (7) \\
 & & \\
\text{BC:} & \left.\begin{array}{l} u_0(n), \ldots, u_{\tilde{M}-1}(n) = 0 \\ u_{K+\tilde{M}}(n), \ldots, u_{K+M-1}(n) = 0 \end{array}\right\} ; \qquad n \in \bar{T} & (8)
\end{array}
\right.
$$

<u>Proof:</u>

Let us use a simplified notation:

$$u_k(n) = u(x_k, t_n)$$

$$z_k(n) = z(x_k, t_n)$$

$$w_k(n) = w_{x_k}(t_n)$$

$$v(n) = v(t_n)$$

$$\beta_k = \beta(x_k)$$

a) The discrete observation in (6) comes from (2) naturally, for all interior points (x_k, t_n) such that $(n,k) \in S \times T$.

b) Use forward and backward operators (see last chapter, section 3.4)

$$\frac{1}{\delta_x^m} \Delta_{\delta_x}^m \quad , \quad \frac{1}{\delta_x^m} \Delta_{\delta_x}^{-m} \quad \text{and} \quad \frac{1}{\delta_t^m} \Delta_{\delta_t}^m$$

to approximate partial derivatives at boundary points $x = x_0 = 0$,

$x = x_{K+M-1} = \ell$ and $t = t_0 = 0$, respectively. So (7) and (8) are approximations for (3) and (4).

c) Use centered operators

$$D^m_{\delta_x} \quad \text{and} \quad D^m_{\delta_t}$$

to approximate partial derivatives at interior points (x_k, t_n) such that $(k,n) \in S \times T'$. In this way, as shown in the second part of the preceding chapter, we have (5) from (1). Where the coefficients

$$a^m_k = a_m(x_k) \quad \text{and} \quad c^m_k = c_m(x_k)$$

are such as introduced on lemma (L - 3.2). □

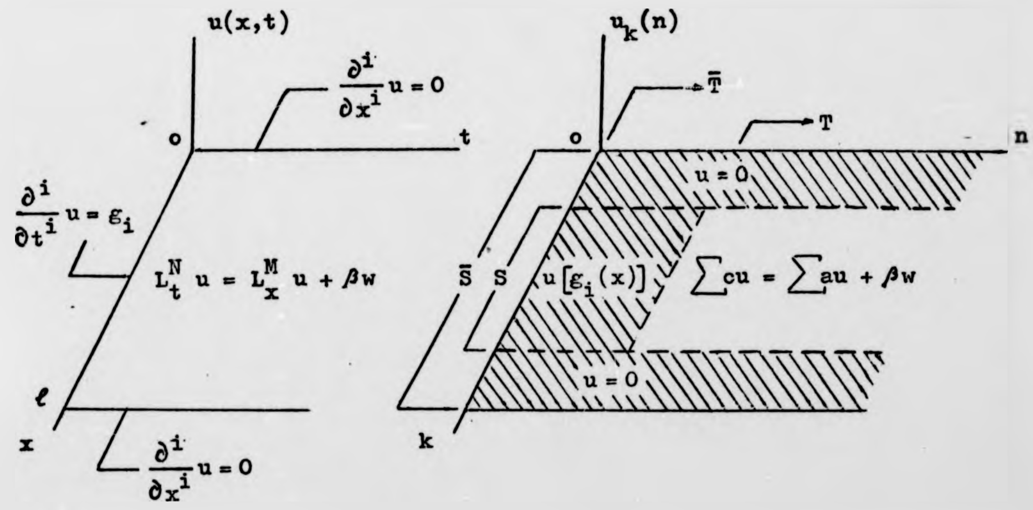

Fig. 6: Continuous and Discrete Formulations.

Remarks:

1) By lemma (L - 3.2) we have the following results (note that $\alpha_0(x) = 0$):

$$a_k^{i+\tilde{M}+1} = \sum_{m=m_\alpha(i)}^{M} (-1)^{\bar{m}-i} \binom{m}{i+\tilde{m}} \frac{\alpha_m(x_k)}{\delta_x^m} \tag{9}$$

for each $i = -\tilde{M}, -\tilde{M}+1, \ldots, \bar{M}$ and $k \in S$; where:

$$m_\alpha(i) = \begin{cases} -2i & ; \quad \text{if } i \le -1 \\ 1 & ; \quad \text{if } i = 0 \\ 2i-1 & ; \quad \text{if } i \ge 1 \end{cases} \quad .$$

$$c_k^{j+\tilde{N}+1} = \sum_{m=m_\gamma(j)}^{N} (-1)^{\bar{m}-j} \binom{m}{j+\tilde{m}} \frac{\gamma_m(x_k)}{\delta_t^m} \tag{10}$$

for each $j = -\tilde{N}, -\tilde{N}+1, \ldots, \bar{N}$ and $k \in S$; where:

$$m_\gamma(j) = \begin{cases} -2j & ; \quad \text{if } j = -1 \\ 0 & ; \quad \text{if } j = 0 \\ 2j-1 & ; \quad \text{if } j = 1 \end{cases} \quad .$$

$$\sum_{m=1}^{M+1} a_k^m = 0 \; ; \qquad \forall \; k \in S \quad . \tag{11}$$

$$\sum_{m=1}^{N+1} c_k^m = \gamma_0(x_k) \; ; \qquad \forall \; k \in S \quad . \tag{12}$$

2) In particular, from (10), the coefficient c_k^{N+1} is such that:

$$c_k^{N+1} = \begin{cases} \dfrac{1}{\delta_t^N} \gamma_N(x_k) & ; \quad \text{if } N \in Z_o \\ \dfrac{1}{\delta_t^N} \left[\gamma_N(x_k) + \delta_t \, \gamma_{N-1}(x_k)\right] & ; \quad \text{if } N \in Z_e \end{cases} \quad .$$

Remarks:

1) By lemma (L - 3.2) we have the following results (note that $\alpha_0(x) = 0$):

$$a_k^{i+\tilde{M}+1} = \sum_{m=m_\alpha(i)}^{M} (-1)^{\bar{m}-i} \binom{m}{i+\tilde{m}} \frac{\alpha_m(x_k)}{\delta_x^m} \tag{9}$$

for each $i = -\tilde{M}, -\tilde{M}+1, \ldots, \bar{M}$ and $k \in S$; where:

$$m_\alpha(i) = \begin{cases} -2i \ ; & \text{if } \ i \le -1 \\ 1 \ ; & \text{if } \ i = 0 \\ 2i-1 \ ; & \text{if } \ i \ge 1 \end{cases} .$$

$$c_k^{j+\tilde{N}+1} = \sum_{m=m_\gamma(j)}^{N} (-1)^{\bar{m}-j} \binom{m}{j+\tilde{m}} \frac{\gamma_m(x_k)}{\delta_t^m} \tag{10}$$

for each $j = -\tilde{N}, -\tilde{N}+1, \ldots, \bar{N}$ and $k \in S$; where:

$$m_\gamma(j) = \begin{cases} -2j \ ; & \text{if } \ j = -1 \\ 0 \ ; & \text{if } \ j = 0 \\ 2j-1 \ ; & \text{if } \ j = 1 \end{cases} .$$

$$\sum_{m=1}^{M+1} a_k^m = 0 \ ; \qquad \forall \ k \in S \ . \tag{11}$$

$$\sum_{m=1}^{N+1} c_k^m = \gamma_0(x_k) \ ; \qquad \forall \ k \in S \ . \tag{12}$$

2) In particular, from (10), the coefficient c_k^{N+1} is such that:

$$c_k^{N+1} = \begin{cases} \dfrac{1}{\delta_t^N} \gamma_N(x_k) \ ; & \text{if } \ N \in Z_o \\ \dfrac{1}{\delta_t^N} \left[\gamma_N(x_k) + \delta_t \, \gamma_{N-1}(x_k) \right] \ ; & \text{if } \ N \in Z_e \end{cases} .$$

But we have already assumed that $\gamma_N(x) \neq 0$ for all $x \in X$ and $\delta_t \gamma_{N-1}(x) = \neq -\gamma_N(n)$ for all $x \in X$ if $n \in Z_e$.

Then:

$$c_k^{N+1} \neq 0 \; ; \quad \forall\, k \in S \; .$$

3) Since $\beta(x) \neq 0$ for all $x \in X$, we have:

$$\beta_k \neq 0 \; ; \quad \forall\, k \in S \; .$$

Equivalent discrete-time LPS: Let R^k denote the k-dimensional Euclidian space, $BL(R^k)$ the normed linear space of all bounded linear operators from R^k into itself, and define[6]:

$$\left.\begin{array}{ll} \underline{u}(n) = (u_{\tilde{M}}(n),\ldots,u_{K+\tilde{M}-1}(n)) \; ; & \forall\, n \in \bar{T} \\ \underline{w}(n) = (w_{\tilde{M}}(n),\ldots,w_{K+\tilde{M}-1}(n)) \; ; & \forall\, n \in T' \\ \underline{z}(n) = (z_{\tilde{M}}(n),\ldots,z_{K+\tilde{M}-1}(n)) \; ; & \forall\, n \in T \end{array}\right\} \text{random vectors in } R^K$$

$$\underline{d} = d(1,1,\ldots,1) \quad \in R^K$$

$$C_m = \begin{bmatrix} c^m_{\tilde{M}} & & \\ & \ddots & \\ & & c^m_{K+\tilde{M}-1} \end{bmatrix} \in BL(R^K) \; ; \quad \forall\, m = 1,2,\ldots,N+1$$

$$B = \begin{bmatrix} \beta_{\tilde{M}} & & \\ & \ddots & \\ & & \beta_{K+\tilde{M}-1} \end{bmatrix} \in BL(R^K)$$

---

6 The quantities in R^k and $BL(R^k)$ are represented with respect to the standard basis in R^k (i.e., $\{\underline{e}_i = (0,\ldots,1,\ldots,0) \; ; \; i = 1,\ldots,k\}$ with a 1 in the i th place and zeros elsewhere).

$$
A = \begin{bmatrix}
a_{\tilde{\tilde{M}}}^{\tilde{\tilde{M}}+1} & \cdots & a_{\tilde{M}}^{M+1} & & & \\
\vdots & & & \ddots & & \\
a_{2\tilde{\tilde{M}}}^{1} & \cdots\cdots & & a_{2\tilde{M}}^{M+1} & & \\
& \ddots & & \cdots\cdots & \ddots & \\
& & a_{K+\tilde{\tilde{M}}-\bar{M}-1}^{1} & \cdots\cdots & & a_{K+\tilde{M}-\bar{M}-1}^{M+1} \\
& & & \ddots & & \vdots \\
& & & a_{K+\tilde{\tilde{M}}-1}^{1} & \cdots & a_{K+\tilde{\tilde{M}}-1}^{\tilde{M}+1}
\end{bmatrix} \in BL(R^K)
$$

The approximate discrete version (5) - (8) can be written as a vector difference equation (discrete-time LPS), as follows:

From (5) and (8) (with $n \in T'$) we get

$$\sum_{m=1}^{N+1} C_m \, \underline{u}(n+m-\tilde{\tilde{N}}-1) = A\,\underline{u}(n) + B\,\underline{w}(n)$$

$$\therefore \qquad C_{N+1}\, \underline{u}(n+\bar{N}) = -\sum_{m=1}^{N} C_m \, \underline{u}(n+m-\tilde{\tilde{N}}-1) + A\,\underline{u}(n) + B\,\underline{w}(n) \; .$$

Since $\quad c_k^{N+1} \neq 0 \quad \forall\, k \in S \;, \quad \exists\, C_{N+1}^{-1}$. Hence:

$$\underline{\underline{u}}(n+\bar{N}) = \sum_{m=1}^{N} A_m \, \underline{u}(n+m-\tilde{\tilde{N}}-1) + B_N \, \underline{w}(n) \; ; \qquad n \in T'$$

where $\{A_m \in BL(R^K) \; ; \;\; m = 1,\ldots,N\}$ and $B_N \in BL(R^K)$ are such that

$$A_m = -C_{N+1}^{-1}\, C_m \; ; \qquad m \neq \tilde{N}+1$$

$$A_{\tilde{N}+1} = C_{N+1}^{-1}\,(A - C_{\tilde{N}+1})$$

$$B_N = C_{N+1}^{-1}\, B$$

Adding initial conditions supplied by (7)

$$\underline{u}(0),\ldots,\underline{u}(N-1)$$

we get a full process description for all $n \in \overline{T}$, with observations given by (5)

$$\underline{z}(n) = h\,\underline{u}(n) + \underline{d}\,v(n) \; ; \qquad n \in T \; .$$

Now define $\{\underline{y}_m(n) \; ; \; m=1,\ldots,N\}$, sets of N random vectors in R^K for each $n \in T'$, as follows:

$$\underline{y}_m(n) = \underline{u}(n+m-\tilde{N}-1) \; ; \qquad n \in T' \; .$$

Then:

$$\underline{y}_m(\tilde{N}) = \underline{u}(m-1) \qquad \text{(initial state)}$$

$$\underline{y}_m(n+1) = \underline{y}_{m+1}(n) \; ; \qquad m \neq N$$

$$\underline{y}_N(n+1) = \underline{u}(n+\bar{N}) = \sum_{m=1}^{N} A_m\, \underline{y}_m(n) + B_N\, \underline{w}(n)$$

$$\underline{y}_{\tilde{N}+1}(n) = \underline{u}(n) \qquad \text{(output)}$$

In this way we get the following model of a linear discrete-time LPS, which is equivalent to the description given in (5) - (8):

$$\begin{cases} \text{MODEL:} & \underline{y}(n+1) = F\,\underline{y}(n) + G\,\underline{w}(n)\ ; \qquad \underline{y}(\tilde{N}) \text{ given} & (13) \\ \text{OBS:} & \underline{z}(n) = H\,\underline{y}(n) + \underline{d}\,v(n)\ ; \qquad n \in T & (14) \end{cases}$$

where:

i) $\underline{y}(n) = (\underline{y}_1(n), \ldots, \underline{y}_N(n))$; $n \in T'$: random vectors in R^{NxK}

ii) $$F = \left[\begin{array}{c|ccc} 0 & I & & \\ \vdots & & \ddots & \\ 0 & & & I \\ \hline A_1 \cdots & A_{\tilde{N}+1} & \cdots & A_N \end{array}\right] \in BL(R^{NxK})$$

iii) $$G = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ B_N \end{bmatrix} \quad : R^K \longrightarrow R^{NxK}$$

iv) $H = h\,[0 \ldots 0\ I\ 0 \ldots 0] \quad : R^{NxK} \longrightarrow R^K$

The identity and null operators, I and 0, are in $BL(R^K)$. In case of H, I is placed at $(\tilde{N}+1)$ th position.

Remarks:

1) (Nonhomogeneous boundary conditions). If we assume (4') we get

$$\text{BC:} \quad \begin{matrix} u_0(n), \ldots, u_{\tilde{N}-1}(n) & \text{given by } h_i^o(x_k) \\ u_{K+\tilde{M}}(n), \ldots, u_{K+M-1}(n) & \text{given by } h_i^\ell(x_k) \end{matrix} \quad ; \quad n \in \bar{T}$$

instead of (8), and the model in (13) becomes:

$$\underline{y}(n+1) = F\,\underline{y}(n) + G\,\underline{w}(n) + G_r\,\underline{u}_r(n)\ ; \qquad \underline{y}(\tilde{N}) \text{ given}$$

where:

i) $\underline{u}_\Gamma(n) = (u_0(n),\dots,u_{\tilde{M}-1}(n),u_{K+\tilde{M}}(n),\dots,u_{K+M-1}(n))$, in R^M, represents the action of the boundary functions h_i^o and h_i^ℓ. It can be thought as an input (or disturbance) vector operating at the spatial boundary of the discrete version.

ii) $$G_\Gamma = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ B_\Gamma \end{bmatrix} \quad : \ R^M \longrightarrow R^{NxK}$$

iii) $$B_\Gamma = C_{N+1}^{-1} A_\Gamma \qquad : \ R^M \longrightarrow R^K$$

iv) $$A_\Gamma = \left[\begin{array}{ccc|ccc} a_{\tilde{M}}^1 & \cdots & a_{\tilde{M}}^{\tilde{M}} & & & \\ & \ddots & \vdots & & & \\ & & a_{2\tilde{M}-1}^1 & & & \\ & & & & & \\ & & & a_{K+\tilde{M}-\tilde{M}}^{M+1} & & \\ & & & \vdots & \ddots & \\ & & & a_{K+\tilde{M}-1}^{\tilde{M}+2} & \cdots & a_{K+\tilde{M}-1}^{M+1} \end{array}\right] \quad : \ R^M \longrightarrow R^K$$

Note that the coefficients of A_Γ are those $\{a_k^m, k\in S, m=1,\dots,M+1\}$, which are not coefficients of A. This can be easily checked by writing down the approximation of L_x^M (i.e.: $\sum_{m=1}^{M+1} a_k^m u_{k+m-\tilde{M}-1}(n)$) $\forall k\in S$.

2) (Lower order models). It is quite obvious that the approximating model (5) (and so (13)) will be more accurate for cases of lower order

where:

i) $\underline{u}_\Gamma(n) = (u_0(n),\dots,u_{\tilde{M}-1}(n),u_{K+\tilde{M}}(n),\dots,u_{K+M-1}(n))$, in R^M, represents the action of the boundary functions h_i^0 and h_i^ℓ. It can be thought as an input (or disturbance) vector operating at the spatial boundary of the discrete version.

ii) $$G_\Gamma = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ B_\Gamma \end{bmatrix} \quad : \; R^M \to R^{N \times K}$$

iii) $$B_\Gamma = C_{N+1}^{-1} A_\Gamma \qquad : \; R^M \to R^K$$

iv) $$A_\Gamma = \left[\begin{array}{ccc:ccc} a^1_{\tilde{M}} & \cdots & a^{\tilde{M}}_{\tilde{M}} & & & \\ & \ddots & \vdots & & & \\ & & a^1_{2\tilde{M}-1} & & & \\ & & & a^{M+1}_{K+\tilde{M}-M} & & \\ & & & \vdots & \ddots & \\ & & & a^{\tilde{M}+2}_{K+\tilde{M}-1} & \cdots & a^{M+1}_{K+\tilde{M}-1} \end{array}\right] \quad : \; R^M \to R^K$$

Note that the coefficients of A_Γ are those $\{a_k^m,\ k \in S,\ m = 1,\dots,M+1\}$, which are not coefficients of A. This can be easily checked by writing down the approximation of L_x^M (i.e.: $\sum_{m=1}^{M+1} a_k^m\, u_{k+m-\tilde{M}-1}(n)$) $\forall k \in S$.

2) (Lower order models). It is quite obvious that the approximating model (5) (and so (13)) will be more accurate for cases of lower order

as commented in chapter 3. For $M \leq 2$ and $N \leq 2$ (what represents the majority of cases of practical and theoretical interest), our general approximation procedure is reduced to the classical finite-difference technique (see [4] for a particular case with $M=2$ and $N=1$).

<u>A basic observation equation</u>: Finally we present a relation expressing the observation dynamics, which is a fundamental step towards the identification procedure introduced in next sections.

Let $\{\mathcal{T}_n \,;\; n \in \mathring{T}\}$ denote a class of finite sets as follows:

$$\mathcal{T}_n = \{n-\tilde{N}, n-\tilde{M}+1, \dots, n+\tilde{N}\} \subset T$$

and, for notational simplicity, define:

i) $$z_k(\mathcal{T}_n) = \sum_{m=1}^{N+1} c_k^m \, z_k(n+m-\tilde{N}-1)$$

ii) $$\nu_k(\mathcal{T}_n) = d \sum_{m=1}^{M+1} c_k^m \, v(n+m-\tilde{N}-1) + h\beta_k \, w_k(n)$$

iii) $$\underline{a}_k = (a_k^1, \dots, a_k^{M+1}) \in R^{M+1}$$

iv) $\underline{z}_k(n) = (z_{k-\tilde{M}}(n), \dots, z_{k+\tilde{M}}(n))$: random vectors in R^{M+1}.

With $\langle \,;\, \rangle$ standing for the inner product in R^{M+1}, we claim that:

<u>Proposition (P - 4.1)</u>: For all $(k,n) \in \mathring{S} \times \mathring{T}$,

$$z_k(\mathcal{T}_n) = \langle \underline{a}_k \,;\, \underline{z}_k(n) \rangle + \nu_k(\mathcal{T}_n)$$

Proof:

Since $h \neq 0$, we get from (6) and (7)

$$\sum_{m=1}^{N+1} c_k^m \, z_k(n+m-\tilde{N}-1) = \sum_{m=1}^{M+1} a_k^m \, z_{k+m-\tilde{M}-1}(n) +$$

$$+ d\left[\sum_{m=1}^{N+1} c_k^m \, v(n+m-\tilde{N}-1) - v(n)\sum_{m=1}^{M+1} a_k^m\right] + h\,\beta_k \, w_k(n)$$

for all (k,n) such that $k-\tilde{M},\ldots,k+\bar{M} \in S$ and $n-\tilde{N},\ldots,n+\bar{N} \in T$, which means: $(k,n) \in \overset{o}{S} \times \overset{o}{T}$. But

$$\sum_{m=1}^{M+1} a_k^m = 0 \; ; \qquad \forall\, k \in S \; .$$

Hence we have $\nu_k(\mathcal{Z}_n)$, independent of $\underline{a}_k$, as defined before. $\square$

4.3 - PARAMETERS IN EXPLICIT FORM

Our goal in this section is to deduce from (P - 4.1) a relation between the parameter vector

$$\underline{a}_k = (a_k^1, \ldots, a_k^{M+1}) \quad \in \; R^{M+1}$$

and the observations $z_k(n)$, which is suitable for applying recursive identification algorithms. First, we introduce some notation:

Notation:

1) A pair of single bars, $|\;|$, stands for the absolute value (modulus) of a scalar quantity.

2) A pair of double bars, $\| \ \|$, denotes the standard Euclidian norm in R^k, as well as the induced uniform norm of an operator (or its matrix) in $BL(R^k)$. That is:

$$\|A\| = \sup_{\underline{x} \neq 0} \frac{\|A\,\underline{x}\|}{\|\underline{x}\|} = \max_{\|\underline{x}\|=1} \|A\,\underline{x}\|$$

for $\underline{x} \in R^k$, $A \in BL(R^k)$ and $\|\underline{x}\|^2 = \langle \underline{x} \,;\, \underline{x} \rangle$.

3) A star, *, denotes the transpose of a matrix in the usual way; and the transpose of a vector when it is (notationally) written as a column vector. In this way we have:

$$\underline{x}\,\underline{y}^* \ : \quad R^k \rightarrow R^k$$

$$\underline{x}^*\,\underline{y} = \langle \underline{x} \,;\, \underline{y} \rangle$$

for any $\underline{x}$ and $\underline{y}$ in R^k.

4) The symbols, $E\{\ \}$ and $Cov\{\ ;\ \}$, stand for the expectation and covariance operators, respectively.

5) The Kronecker delta function $\delta(i)$ is defined by

$$\delta(i) = \begin{cases} 1 \,; & \text{if } i = 0 \\ 0 \,; & \text{otherwise} \end{cases}$$

Now, make the following assumptions on the linear discrete-time system described in (13) and (14):

Assumption (A - 4.1): (Stability) The space and time sampling rates, δ_x and δ_t, are choosen such that the system in (13) is stable in the following sense: There exist constants K_o and ρ $(0 < \rho < 1)$ such that

$$\|F\|^n < K_o\,\rho^n \,; \qquad \forall n \in \bar{T}$$

Assumption (A - 4.2): (Stationarity) $\{v(n)\ ;\ n \in T\}$ and $\{\underline{w}(n)\ ;\ n \in T'\}$ are real and R^K-valued second-order wide-sense stationary random sequences[7], respectively, with the following additional conditions:

i) $E\{\underline{w}(n)\} = \underline{0}$

ii) $\mathrm{Cov}\{\underline{w}(i)\ ;\ \underline{w}(j)\} = E\{\underline{w}(i)\ \underline{w}^*(j)\} = C_w\,\delta(i-j)$

iii) $E\{v(n)\} = \eta_v$

iv) $\mathrm{Cov}\{v(i)\ ;\ v(j)\} = E\{v(i)\ v(j)\} - \eta_v^2 = \sigma_v(|i-j|)$

v) $\mathrm{Cov}\{\underline{d}\,v(i)\ ;\ \underline{w}(i)\} = \underline{d}\,E\{v(i)\ \underline{w}^*(j)\} = 0$

vi) $\underline{w}(n)$ and $v(n)$ have finite moments up to the 4th order.

Where C_w is a symmetric positive definite matrix in $BL(R^K)$.

Assumption (A - 4.3): (Steady state) The initial state response is assumed to have died out before identification begins. Since the input disturbance $\underline{w}(n)$ and the observation noise $v(n)$ are wide-sense stationary and the system under consideration is time-invariant, this assumption implies that the state $\underline{y}(n)$ and observation $\underline{z}(n)$ processes will also be wide-sense stationary during the identification procedure (see, for example, [5]-[7]).

Assumption (A - 4.4): (Finite transient time) Let $N_t \geq N+\tilde{N} \in \mathring{T}$ be a sufficient large integer such that the steady state assumption at the time $n+N_t$ is valid for all $n \in \overline{T}$.

---

7 The term "random sequence" will be used to denote a "discrete-time stochastic process".

Finally, for each $k \in \overset{o}{S}$, define:

i) $\varepsilon_k = -d^2 \sum_{m=1}^{N+1} c_k^m \left[\sigma_v(m-\tilde{N}-1) + \eta_v^2\right]$ [8]

ii) $\underline{\varepsilon}_k = \varepsilon_k(1,\dots,1) \quad \in \quad R^{M+1}$

iii) $\underline{q}_k = E\{\underline{z}(n+N_t)\ z_k(\tau_{n+N_t})\} + \underline{\varepsilon}_k \quad \in \quad R^{M+1}$

iv) $Q_k^{-1} = E\{\underline{z}_k(n+N_t)\ \underline{z}_k^*(n+N_t)\} \quad \in \quad BL(R^{M+1})$

Proposition (P - 4.2): For all k in $\overset{o}{S}$ and n in $\bar{T}$,

$$E\{\underline{z}_k(n+N_t)\ \nu_k(\tau_{n+N_t})\} = -\underline{\varepsilon}_k$$

Proposition (P - 4.3): There exists the symmetric positive definite matrix $Q_k = [Q_k^{-1}]^{-1} \in BL(R^{M+1})$, for all $k \in \overset{o}{S}$.

Proposition (P - 4.4): Both $\underline{q}_k$ and Q_k do not depend on n.

---

8 By (12) we get :

$$\varepsilon_k = -d^2\left[\eta_v^2\ \gamma_o(x_k) + \sum_{m=1}^{N+1} c_k^m\ \sigma_v(|m-\tilde{N}-1|)\right]$$

If $\sigma_v(|i-j|) = \sigma_v^2\ \delta(i-j)$, $\sigma_v^2 > 0$, we have from (10):

$$\varepsilon_k = -d^2\left[\eta_v^2\ \gamma_o(x_k) + \sigma_v^2 \sum_{m=0}^{N} (-1)^{\bar{m}} \binom{m}{\tilde{m}} \frac{\gamma_m(x_k)}{\delta_t^m}\right]$$

Note also that in [4] ε_k is defined slightly different. There, it is divided by $c_k^{N+1} = \delta_t^{-1}$ for the particular case of N=1, $\gamma_o = -\alpha_o$ and $\gamma_1 = 1$ (see section 4.6).

Proof:

(P - 4.2): Using finite induction we get the solution of the equation (13):

$$\underline{y}(n+\tilde{\tilde{N}}+1) = F^{n+1}\,\underline{y}(\tilde{\tilde{N}}) + \sum_{m=0}^{n} F^{n-m}\,G\,\underline{w}(m+\tilde{\tilde{N}}) \;; \qquad n \leq \bar{\bar{T}} .$$

By assumption (A - 4.3) and (A - 4.4), the first term in the right hand side is supposed to have died out for any time $n \geq N_t$.[9] Thus,

$$\underline{y}(n+N_t) = \sum_{m=0}^{n+N_t-\tilde{\tilde{N}}-1} F^{n+N_t-m-\tilde{\tilde{N}}-1}\,G\,\underline{w}(m+\tilde{\tilde{N}})$$

and from (14) we get

$$\underline{z}(n+N_t) = H\,\underline{y}(n+N_t) + \underline{d}\,v(n+N_t) .$$

Now, with

$$\nu(\mathcal{Z}_{n+N_t}) = d \sum_{m=1}^{N+1} c_k^m\, v(n+N_t+m-\tilde{\tilde{N}}-1) + h\,\beta_k\, w_k(n+N_t),$$

assumption (A - 4.2), and using the linearity of the expectation operator [1], we have:

$$E\{\underline{z}(n+N_t)\,\nu(\mathcal{Z}_{n+N_t})\} = d\,\underline{d} \sum_{m=1}^{N+1} c_k^m \left[\delta_v(|m-\tilde{\tilde{N}}-1|) + \eta_v^2\right]$$

$$= -\,\varepsilon_k(1,\ldots,1) \quad \in R^K .$$

[9] Actually, the assumptions (A - 4.3) and (A - 4.4) can be thought as a consequence of (A - 4.1) when one is considering the asymptotic behaviour. That is:

$$0 \leq \|F^n\| \leq \|F\|^n \leq K_o\,\rho^n \;; \qquad 0<\rho<1$$

$$\therefore \quad \lim_{n\to\infty} \|F^n\| \leq K_o \lim_{n\to\infty} \rho^n = 0 \quad \Rightarrow \quad F^n \to 0 \quad \text{as} \quad n \to \infty$$

Finally, define:

$$J_k = \left[\, \emptyset \,\vdots\, I \,\vdots\, \emptyset \,\right] \qquad : \; R^K \longrightarrow R^{M+1}$$

where there are $k-2\bar{N}$ zero columns before $I \in BL(R^{M+1})$ and, of course, $K+\tilde{N}-\bar{N}-k-1$ zero columns after I.

So:

$$\underline{z}_k(n) = J_k\, \underline{z}(n) \; ; \qquad \forall\, k \in \overset{o}{S} .$$

Hence:

$$E\{\underline{z}_k(n)\, \nu(\mathcal{Z}_{n+N_t})\} = J_k\, E\{\underline{z}(n)\, \nu(\mathcal{Z}_{n+N_t})\}$$

and (P – 4.2) is proved.

(P – 4.3): It is easy to show that: [10]

$$\underline{z}(n+\bar{N}) = H F^{\bar{N}}\, \underline{y}(n) + H F^{\bar{N}-1} G\, \underline{w}(n) + \underline{d}\, v(n+\bar{N}) \; ; \qquad n \in T'$$

and

$$H F^{\bar{N}-1} = h\left[0 \;\dots\; 0 \; I\right] \qquad : \; R^{NxK} \longrightarrow R^K$$

$$H F^{\bar{N}} = h\left[A_1 \;\dots\; A_N\right] \qquad : \; R^{NxK} \longrightarrow R^K$$

Now define:

$$\Phi = H F^{\bar{N}} \qquad : \; R^{NxK} \longrightarrow R^K$$

$$\Gamma = H F^{\bar{N}-1} G = h B_N \qquad \in \; BL(R^K)$$

$$C_y = \mathrm{Cov}\{\underline{y}(n+N_t) \; ; \; \underline{y}(n+N_t)\} \qquad \in \; BL(R^{NxK})$$

$$C_z = \mathrm{Cov}\{\underline{z}(n+N_t) \; ; \; \underline{z}(n+N_t)\} \qquad \in \; BL(R^K)$$

---

10 This result comes from equations (13) and (14) or, equivalently, from the solution for $\underline{y}(n)$, $n > \tilde{N}$, as introduced at the begining of this proof. Note that $H F^m G = 0$, $\forall\, m = 0,\dots,\bar{N}-2$.

Then

$$\underline{z}(n+N_t+\bar{N}) = \Phi\underline{y}(n+N_t) + \Gamma\underline{w}(n+N_t) + \underline{d}\,v(n+N_t+\bar{N}) \; ; \qquad n \in T' \; .$$

From (A - 4.3) and (A - 4.4), C_y and C_z do not depend on n. So, by (A - 4.2), we get:[11]

$$C_z = \Phi\, C_y\, \Phi^* + \Gamma\, C_w\, \Gamma^* + \delta_v(0)\, \underline{d}\, \underline{d}^*$$

and $Q_k^{-1} \in BL(R^{N+1})$ is given in terms of $C_z \in BL(R^K)$ as follows:

$$\underline{z}_k(n) = J_k\, \underline{z}(n) \quad \rightarrow \quad Q_k^{-1} = J_k\, C_z\, J_k^* \; .$$

Since $C_w > 0$ (positive definite) $\in BL(R^K)$, and all eigenvalues of

$$\Gamma = h\, B_N = -h\, C_{N+1}^{-1}\, B \qquad \in \quad BL(R^K)$$

are different from zero[12],

$$\Gamma\, C_w\, \Gamma^* > 0 \quad \rightarrow \quad C_z > 0 \; .$$

But, if $C_z > 0$, all principal minors of C_z are positive. In particular, all successive principal minors of the symmetric matrix

$$Q_k^{-1} = J_k\, C_z\, J_k^* \qquad \in \quad BL(R^{N+1})$$

---

11 Note that:

$$Cov\{\Phi\, \underline{y}(n) \; ; \; \underline{w}(n)\} = 0 \; ; \qquad \forall\, n \in T'$$

$$E\{\underline{w}(n+N_t)\} = \underline{0} \rightarrow E\{\underline{y}(n+N_t)\} = \underline{0} \rightarrow E\{\underline{z}(n+N_t)\} = \eta_v\, \underline{d} \; ; \quad \forall\, n \in T$$

12 Recall: $0 \neq |\beta_k| < \infty$ and $0 \neq |C_k^{N+1}| < \infty$; $\qquad \forall\, k \in S$.

are positive. Hence $Q_k^{-1} > 0$ (Sylvester's criteria. See, for example, [8] pp. 306).[13] So, there exists the symmetric

$$Q_k = \left[Q_k^{-1}\right]^{-1} > 0 \quad \in \quad BL(R^{M+1}) \; ; \quad \forall k \in \overset{\circ}{S} \; .$$

(P - 4.4): This result comes directly from (A - 4.2) - (A - 4.4): wide-sense stationary random sequences (for details see, for example, [5]-[7]).

□

<u>Remark</u>: (Random initial state) Assume that all eigenvalues of the system matrix F in $BL(R^{NxK})$ are different from zero (or equivalently, $\det(F) \neq 0$), and define

$$\mathrm{Cov}\{\underline{y}(n) \; ; \; \underline{y}(n)\} = \begin{cases} C_y(n) \; ; & \text{if } \ \ddot{N} \leq n < N_t \\ C_y \; ; & \text{if } \ n \geq N_t \end{cases}$$

It is well known (e.g., see [5]-[7]) that

$$C_y(n+1) = F\, C_y(n)\, F^* + G\, C_w\, G^*$$

for all $n \in T'$. So, if the initial state

$$\underline{y}(\ddot{N}) = (\underline{u}(0), \ldots, \underline{u}(N-1))$$

is a random vector in R^{NxK} such that

$$C_y(\ddot{N}) > 0$$

---

13 Note that, in case of $(M+K) \in Z_o$,

$$Q_{k_o}^{-1} = J_{k_o}\; C_z\; J_{k_o}^* \quad \in \quad BL(R^{M+1})$$

is the $\overline{(M+1)}$th "inner" of $C_z \in BL(R^K)$ [9], where $k_o = (K+M-1)/2$.

in $BL(R^{NxK})$, we get

$$C_y(n) > 0 \; ; \qquad \forall n \in T'$$

in $BL(R^{NxK})$, and so:

$$\Phi \; C_y \; \Phi^* > 0 \quad \rightarrow \quad C_z > 0$$

in $BL(R^K)$. Hence, in this case, the conditions

$$\beta(x) \neq 0 \; ; \qquad \forall x \in X$$

and

$$C_w > 0 \quad \text{in} \quad BL(R^K)$$

can be omited, since they are imposed only to ensure that $C_z > 0$ when the initial state is assumed to be deterministic.

<u>Lemma (L - 4.1)</u>: (Explicit Parameter) Let $\underline{q}_k$ and Q_k be as defined before. If the assumptions (A - 4.1) through (A - 4.4) are satisfied, then the parameter vector

$$\underline{a}_k = (a_k^1, \ldots, a_k^{M+1}) \quad \in \quad R^{M+1}$$

introduced in (P - 4.1) can be placed in an explicit form, as follows:

$$\underline{a}_k = Q_k \; \underline{q}_k$$

for all k in $\overset{o}{S}$.

<u>Proof</u>:

By (P - 4.1) through (P - 4.4), we have:

$$z_k(\tau_{n+N_t}) = \underline{z}_k^*(n+N_t) \; \underline{a}_k + \nu_k(\tau_{n+N_t}) \; ; \qquad \forall k \in \overset{o}{S}$$

$\therefore$

$$E\{\underline{z}_k(n+N_t)\, z_k(\tau_{n+N_t})\} = E\{\underline{z}_k(n+N_t)\, \underline{z}_k^*(n+N_t)\}\underline{a}_k + E\{\underline{z}_k(n+N_t)\, \nu_k(\tau_{n+N_t})\} =$$

$$= Q_k^{-1}\, \underline{a}_k - \underline{\xi}_k \, .$$

Then:

$$\underline{a}_k = Q_k\, [E\{\underline{z}_k(n+N_t)\, z_k(\tau_{n+N_t})\} + \underline{\xi}_k] = Q_k\, \underline{g}_k \, ; \qquad \forall k \in \overset{\circ}{S} . \qquad \square$$

4.4 - PARAMETERS IDENTIFICATION

General procedure: Now we consider the identification problem for DPS as formulated in section 4.1, by using the reduction to a finite-dimensional state space introduced in section 4.2. In this way, our identification procedure comprises two basic steps:

1st) By using noisy observations $\{z_k(n+N_t) \; ; \; n \in \overline{T}\}$ available in each $k \in S$, determine the coefficients $\{a_k^m \; ; \; m = 1,\ldots,M+1\}$ appearing in the discrete version of the spatial-differential operator L_x^M, as in (5). Since the observation process in R^{M+1}

$$\underline{z}_k(n+N_t) = (z_{k-\overline{M}}(n+N_t),\ldots,z_{k+\overline{M}}(n+N_t))$$

is defined only for $k \in \overset{\circ}{S}$ and

$$\underline{a}_k = (a_k^1,\ldots,a_k^{M+1}) \quad \in \quad R^{M+1}$$

is related to $\underline{z}_k(n+N_t)$ as in (L - 4.1), this first step can be briefly stated as follows: "given $\{z_k(n+N_t) \; ; \; n \in \overline{T}\} \; \forall k \in S$, determine $\underline{a}_k \; \forall k \in \overset{\circ}{S}$".

In (L - 4.1) we got $\underline{a}_k$ in terms of Q_k and $\underline{q}_k$. But it is quite obvious that we cannot use the values of Q_k and $\underline{q}_k$ in order to perform the identification, because both of them depend on the knowledge of the matrix F (and so, they depend on the parameters a_k^m). Therefore, by using the stochastic approximation theory (chapter 3 - Part III) together with the explicit parameter lemma (L - 4.1), we may be able to present an on-line identification algorithm for $\underline{a}_k$, without computing the values of Q_k and $\underline{q}_k$. This first step is the central theme in the remainder of this section.

2nd) On the other hand, we also have the problem of recovering the parameters $\{\alpha_m(x_k)\ ;\ m=1,\dots,M\}$ from $\underline{a}_k$, for each $k \in \overset{o}{S}$. This is a much easier problem than that concerning the first step, and it will be considered in the next section.

Stochastic approximation algorithms: First, by using the results developed in the third part of the preceding chapter, we present an auxiliary lemma for recursive estimation of $\underline{q}_k$ and Q_k^{-1}.[14]

Lemma (L - 4.2): Let $z_k(\tau_{n+N_t})$, $\underline{z}_k(n+N_t)$, $\underline{\varepsilon}_k$, $\underline{q}_k$ and Q_k^{-1} be as defined before and consider the following algorithms in R^{M+1} and $BL(R^{M+1})$, respectively, for each $k \in \overset{o}{S}$:

$$\underline{q}_k(n+1) = [1-\lambda(n)]\,\underline{q}_k(n) + \lambda(n)\left[z_k(\tau_{n+N_t})\,\underline{z}_k(n+N_t) + \underline{\varepsilon}_k\right]\ ; \qquad n \in \bar{T} \quad (15)$$

$$Q_k^{-1}(n+1) = [1-\mu(n)]\,Q_k^{-1}(n) + \mu(n)\left[\underline{z}_k(n+N_t)\,\underline{z}_k^{*}(n+N_t)\right]\ ; \qquad n \in \bar{T} \quad (16)$$

---

14 Note that, by (A - 4.1) and (A - 4.2), $\|\underline{q}_k\| < \infty$ and $\|Q_k^{-1}\| < \infty$ for all $k \in \overset{o}{S}$.

where $\{\lambda(n)\ ;\ \ n\in\bar{T}\}$ and $\{\mu(n)\ ;\ \ n\in\bar{T}\}$ are real sequences, $\underline{q}_k(0)$ is a second-order random vector in R^{M+1} which is independent of $\{\underline{z}_k(n+N_t)\ ;\ n\in\bar{T}\}$ for each $k\in\overset{o}{S}$, and $Q_k^{-1}(0)$ is a second-order random matrix in $BL(R^{M+1})$ which is independent of $\{\underline{z}_k(n+N_t)\ \underline{z}_k^{*}(n+N_t)\ ;\ \ n\in\bar{T}\}$ for each $k\in\overset{o}{S}$.

If:

i) $\lambda(n)\in(0,1)\ ;\quad \sum_{n=0}^{\infty}\lambda(n)=\infty\ ;\quad \sum_{n=0}^{\infty}\lambda^2(n)<\infty$

ii) $\mu(n)\in(0,1)\ ;\quad \sum_{n=0}^{\infty}\mu(n)=\infty\ ;\quad \sum_{n=0}^{\infty}\mu^2(n)<\infty$

Then:

1) $P\{\lim_{n\to\infty}\underline{q}_k(n)=\underline{q}_k\}=1$ and $\lim_{n\to\infty}E\{\|\underline{q}_k(n)-\underline{q}_k\|^2\}=0$

2) $P\{\lim_{n\to\infty}Q_k^{-1}(n)=Q_k^{-1}\}=1$ and $\lim_{n\to\infty}E\{\|Q_k^{-1}(n)-Q_k^{-1}\|^2\}=0$

for each $k\in\overset{o}{S}$.

<u>Proof</u>:

Set, for each $k\in\overset{o}{S}$,

$$y(n)=z_k(\mathcal{Z}_{n+N_t})\ \underline{z}_k(n+N_t)+\underline{\varepsilon}_k\ ;\qquad x_o=\underline{q}_k$$

$$Y(n)=\underline{z}_k(n+N_t)\ \underline{z}_k^{*}(n+N_t)\ ;\qquad X_o=Q_k^{-1}$$

in (C - 3.1) and (C - 3.2), and the proof follows directly by the results of section 3.6 and assumptions (A - 4.1) - (A - 4.4). $\square$

Before introducing the main identification theorem, we need to prove the following propositions:

Proposition (P - 4.5): Let $\{Q_k^{-1}(n)\ ;\ n \in \bar{T}\}$ be a sequence of random matrices in $BL(R^{M+1})$ as defined in (L - 4.2) for each $k \in \overset{\circ}{S}$. If, in addition, $Q_k^{-1}(0)$ is symmetric positive definite, then there exists the symmetric positive definite random matrix

$$Q_k(n) = \left[Q_k^{-1}(n)\right]^{-1}$$

in $BL(R^{M+1})$ for all $n \in \bar{T}$ and $k \in \overset{\circ}{S}$.

Proof:

From (16)

$$Q_k^{-1}(n+1) = [1-\mu(n)]\ Q_k^{-1}(n) + \mu(n)\ \underline{z}_k(n+N_t)\ \underline{z}_k^*(n+N_t)\ .$$

Since $\mu(n) \in (0,1)$ for all $n \in \bar{T}$, $Q_k^{-1}(0)$ is symmetric positive definite for each $k \in \overset{\circ}{S}$, and $\underline{z}_k(n+N_t)\ \underline{z}_k^*(n+N_t)$ is symmetric positive semi-definite for all $(k,n) \in \overset{\circ}{S} \times \bar{T}$; $Q_k^{-1}(n)$ is symmetric positive definite for all $(k,n) \in \overset{\circ}{S} \times \bar{T}$. The existence of $Q_k(n) = \left[Q_k^{-1}(n)\right]^{-1}$, a symmetric positive definite random matrix in $BL(R^{M+1})$, is thus guaranteed for all $(k,n) \in \overset{\circ}{S} \times \bar{T}$. $\square$

Proposition (P - 4.6): Let $A \in BL(R^k)$, $\underline{w} \in R^k$ and assume the existence of A^{-1} and $(A + \underline{w}\,\underline{w}^*)^{-1}$. If $\underline{w}^* A^{-1} \underline{w} \neq -1$, then:[15]

$$(A + \underline{w}\,\underline{w}^*)^{-1} = A^{-1} - (1 + \underline{w}^* A^{-1} \underline{w})^{-1} A^{-1} \underline{w}\,\underline{w}^* A^{-1}\ .$$

15

Note that if $A>0$, then: $\begin{cases} \exists\, A^{-1} > 0 \ \rightarrow\ \underline{w}^* A^{-1} \underline{w} \geq 0 \\ (A + \underline{w}\,\underline{w}^*) > 0 \ \rightarrow\ \exists\, (A + \underline{w}\,\underline{w}^*)^{-1} > 0 \end{cases}$

Proof:

This is a trivial particular case of the "matrix inversion lemma" [6], or "method of modification" [10]. □

Now we can present a stochastic approximation algorithm for identifying the vector $\underline{a}_k$ [16] through noisy observations $\{z_k(n+N_t) ; n \in \bar{T}\}$.

Theorem (T - 4.1): Let I denote the identity matrix in $BL(R^{M+1})$ and $\underline{a}_k \in R^{M+1}$, for each $k \in \overset{o}{S}$, the parameter vector as introduced on proposition (P - 4.1). Also let $z_k(\tau_{n+N_t})$, $\underline{z}_k(n+N_t)$ and $\underline{\varepsilon}_k$ be as defined before, and consider the following algorithm in R^{M+1}, for each $k \in \overset{o}{S}$:

$$(SA-1): \quad \underline{a}_k(n+1) = \frac{1-\lambda(n)}{1-\mu(n)}\left[I - \mu(n)\, Q_k(n+1)\, \underline{z}_k(n+N_t)\, \underline{z}_k^*(n+N_t)\right] \underline{a}_k(n) + $$
$$+ \lambda(n)\, Q_k(n+1)\left[z_k(\tau_{n+N_t})\, \underline{z}_k(n+N_t) + \underline{\varepsilon}_k\right] ; \quad n \in \bar{T}$$

with $Q_k(n)$ taking values in $BL(R^{M+1})$, for each $k \in \overset{o}{S}$, given by

$$(SA-2): \quad Q_k(n+1) =$$
$$= \frac{1}{1-\mu(n)}\left[Q_k(n) - \frac{\mu(n)\, Q_k(n)\, \underline{z}_k(n+N_t)\, \underline{z}_k^*(n+N_t)\, Q_k(n)}{1-\mu(n)+\mu(n)\, \underline{z}_k^*(n+N_t)\, Q_k(n)\, \underline{z}_k(n+N_t)}\right] ; \quad n \in \bar{T}$$

where $\{\lambda(n) ; n \in \bar{T}\}$ and $\{\mu(n) ; n \in \bar{T}\}$ are real sequences, $\underline{a}_k(0)$ is a second-order random vector in R^{M+1} which is independent of $\{\underline{z}_k(n+N_t) ; n \in \bar{T}\}$ for each $k \in \overset{o}{S}$, and $Q_k(0)$ is a second-order random matrix in $BL(R^{M+1})$ which is independent of $\{\underline{z}_k(n+N_t)\, \underline{z}_k^*(n+N_t) ; n \in \bar{T}\}$ for each $k \in \overset{o}{S}$.

---

16 Note: Since $\{\alpha_m(x) \in BV[0,\ell] ; m = 1,\ldots,M\}$, $\|\underline{a}_k\| < \infty$ for all $k \in \overset{o}{S}$, by (9).

If:

i) $\lambda(n) \in (0,1)$; $\sum_{n=0}^{\infty} \lambda(n) = \infty$; $\sum_{n=0}^{\infty} \lambda^2(n) < \infty$

ii) $\mu(n) \in (0,1)$; $\sum_{n=0}^{\infty} \mu(n) = \infty$; $\sum_{n=0}^{\infty} \mu^2(n) < \infty$

iii) $Q_k(0)$ is a symmetric positive definite random matrix in $BL(R^{M+1})$

Then $\underline{a}_k(n)$ converges to $\underline{a}_k$ with probability one for each $k \in \mathring{S}$:

$$P\{\lim_{n\to\infty} \underline{a}_k(n) = \underline{a}_k\} = 1 \; ; \qquad k \in \mathring{S} \; .$$

<u>Proof</u>:

For sake of simplicity we use the following notation:

$$\lambda = \lambda(n)$$

$$\mu = \mu(n)$$

$$z = z_k(\tau_{n+N_t})$$

$$\underline{z} = \underline{z}_k(n+N_t)$$

$$\underline{q} = \underline{q}_k(n)$$

$$\underline{a} = \underline{a}_k(n)$$

$$Q = Q_k(n)$$

$$Q^{-1} = Q_k^{-1}(n)$$

a) First of all let us prove the algorithm (SA-2). From (16)

$$Q_k^{-1}(n+1) = (1-\mu)\, Q^{-1} + \mu \underline{z}\, \underline{z}^* \; .$$

So, by (P - 4.5) and (P - 4.6) with $Q = \left[Q^{-1}\right]^{-1}$, we get

$$Q_k(n+1) = \left[(1-\mu)\, Q^{-1} + \mu \underline{z}\, \underline{z}^*\right]^{-1} =$$

$$= \frac{Q}{1-\mu} - \left(1 - \mu \underline{z}^* \frac{Q}{1-\mu} \underline{z}\right)^{-1} \frac{\mu}{(1-\mu)^2}\, Q\, \underline{z}\, \underline{z}^*\, Q =$$

$$= \frac{1}{1-\mu}\left[Q - \frac{Q\, \underline{z}\, \underline{z}^*\, Q}{1-\mu+\mu \underline{z}^*\, Q\, \underline{z}}\right]$$

since $Q_k(0)$ is symmetric positive definite, and that proves (SA-2). Moreover, note that:

$$Q_k(n+1)\, \underline{z} = \frac{Q}{1-\mu}\left[\frac{(1-\mu)\,\underline{z} + \left[(\underline{z}^*\, Q\, \underline{z})\, \underline{z} - (\underline{z}\, \underline{z}^*\, Q)\, \underline{z}\right]}{1-\mu+\mu\, \underline{z}^*\, Q\, \underline{z}}\right].$$

But

$$(\underline{z}\, \underline{z}^*\, Q)\, \underline{z} = \underline{z}\, (\underline{z}^*\, Q\, \underline{z}) \quad .$$

Hence

$$(\underline{z}^*\, Q\, \underline{z})\, \underline{z} - (\underline{z}\, \underline{z}^*\, Q)\, \underline{z} = (\underline{z}^*\, Q\, \underline{z})\, \underline{z} - \underline{z}\, (\underline{z}^*\, Q\, \underline{z}) = \underline{0} \quad .$$

Then

$$Q_k(n+1)\, \underline{z} = \frac{Q\, \underline{z}}{1-\mu+\mu \underline{z}^*\, Q\, \underline{z}} \tag{17}$$

b) Now define:

$$\underline{a}_k(n) = Q_k(n)\, \underline{q}_k(n) \; ; \qquad n \in \bar{T}$$

for each $k \in \overset{o}{S}$, where $\underline{q}_k(n)$ is given by (15). Thus from (15) and (SA-2), with

$$\omega = (1-\mu) + \mu \underline{z}^*\, Q\, \underline{z}$$

So, by (P - 4.5) and (P - 4.6) with $Q = \left[Q^{-1}\right]^{-1}$, we get

$$Q_k(n+1) = \left[(1-\mu)\, Q^{-1} + \mu\, \underline{z}\, \underline{z}^*\right]^{-1} =$$

$$= \frac{Q}{1-\mu} - (1-\mu\, \underline{z}^* \frac{Q}{1-\mu} \underline{z})^{-1} \frac{\mu}{(1-\mu)^2}\, Q\, \underline{z}\, \underline{z}^*\, Q =$$

$$= \frac{1}{1-\mu}\left[Q - \frac{Q\, \underline{z}\, \underline{z}^*\, Q}{1-\mu+\mu \underline{z}^*\, Q\, \underline{z}}\right]$$

since $Q_k(0)$ is symmetric positive definite, and that proves (SA-2). Moreover, note that:

$$Q_k(n+1)\, \underline{z} = \frac{Q}{1-\mu}\left[\frac{(1-\mu)\, \underline{z} + \left[(\underline{z}^*\, Q\, \underline{z})\, \underline{z} - (\underline{z}\, \underline{z}^*\, Q)\, \underline{z}\right]}{1-\mu+\mu\, \underline{z}^*\, Q\, \underline{z}}\right].$$

But

$$(\underline{z}\, \underline{z}^*\, Q)\, \underline{z} = \underline{z}\, (\underline{z}^*\, Q\, \underline{z}) \quad .$$

Hence

$$(\underline{z}^*\, Q\, \underline{z})\, \underline{z} - (\underline{z}\, \underline{z}^*\, Q)\, \underline{z} = (\underline{z}^*\, Q\, \underline{z})\, \underline{z} - \underline{z}\, (\underline{z}^*\, Q\, \underline{z}) = \underline{0} \quad .$$

Then

$$Q_k(n+1)\, \underline{z} = \frac{Q\, \underline{z}}{1-\mu+\mu\, \underline{z}^*\, Q\, \underline{z}} \tag{17}$$

b) Now define:

$$\underline{a}_k(n) = Q_k(n)\, \underline{q}_k(n) \; ; \qquad n \in \overline{T}$$

for each $k \in \overset{o}{S}$, where $\underline{q}_k(n)$ is given by (15). Thus from (15) and (SA-2), with

$$\omega = (1-\mu) + \mu\, \underline{z}^*\, Q\, \underline{z}$$

we get:

$$\underline{a}_k(n+1) = Q_k(n+1)\, \underline{q}_k(n+1) =$$

$$= Q_k(n+1)\left[(1-\lambda)\,\underline{q} + \lambda z\, \underline{z}\right] + \lambda Q_k(n+1)\, \underline{\varepsilon}_k \; .$$

But

$$Q_k(n+1)\left[(1-\lambda)\,\underline{q} + \lambda z\, \underline{z}\right] =$$

$$= \frac{1}{1-\mu}\left[Q - \frac{\mu Q\, \underline{z}\, \underline{z}^* Q}{\omega}\right]\left[(1-\lambda)\,\underline{q} + \lambda z\, \underline{z}\right] =$$

$$= \frac{1}{1-\mu}\left[(1-\lambda)\, Q\, \underline{q} + \lambda z\, Q\, \underline{z} - \frac{Q\, \underline{z}}{\omega}\left(\mu(1-\lambda)\, \underline{z}^* Q\, \underline{q} + \mu\lambda z\, \underline{z}^* Q\, \underline{z}\right)\right] =$$

$$= \frac{1}{1-\mu}\left[(1-\lambda)\, \underline{a} + \frac{Q\, \underline{z}}{\omega}\left(\lambda z \omega - \mu(1-\lambda)\, \underline{z}^* \underline{a} - \mu\lambda z\, \underline{z}^* Q\, \underline{z}\right)\right] =$$

$$= \frac{1-\lambda}{1-\mu}\, \underline{a} + \frac{Q\, \underline{z}}{\omega}\left[\lambda z\, \frac{\omega - \mu \underline{z}^* Q\, \underline{z}}{1-\mu} - \frac{1-\lambda}{1-\mu}\, \underline{z}^*\, \underline{a}\right]$$

Since

$$\omega - \mu \underline{z}^* Q\, \underline{z} = 1-\mu$$

we have

$$\underline{a}_k(n+1) = \frac{1-\lambda}{1-\mu}\, \underline{a} + \frac{Q\, \underline{z}}{1-\mu+\mu\underline{z}^* Q\, \underline{z}}\left[\lambda z - \frac{1-\lambda}{1-\mu}\, \underline{z}^*\, \underline{a}\right] +$$

$$+ \lambda Q_k(n+1)\, \underline{\varepsilon}_k$$

and so the algorithm (SA-1) comes through (17).

c) The assumptions about $\underline{a}_k(0)$ [17] and $Q_k(0)$ plus (i) - (iii) are sufficient to ensure the convergence in lemma (L - 4.2). So if we define

---

[17] Note that $\underline{q}_k(0) = \left[Q_k(0)\right]^{-1} \underline{a}_k(0)$, by definition of $\{\underline{a}_k(n)\; ;\; n \in \overline{T}\}$.

the events

$$A_k = \left\{ \lim_{n \to \infty} Q_k^{-1}(n) = Q_k^{-1} \right\}$$

$$B_k = \left\{ \lim_{n \to \infty} \underline{q}_k(n) = \underline{q}_k \right\}$$

for each $k \in \overset{o}{S}$, we get by (L - 4.2) that

$$P(A_k) = P(B_k) = 1 \;.$$

Hence

$$P(A_k \cup B_k) = 1$$

$$\therefore \quad P(A_k \cap B_k) = P(A_k) + P(B_k) - P(A_k \cup B_k) = 1 \;.$$

That is

$$P(A_k \cap B_k) = P\left\{ \lim_{n \to \infty} Q_k^{-1}(n) = Q_k^{-1} \;;\; \lim_{n \to \infty} \underline{q}_k(n) = \underline{q}_k \right\} = 1$$

for each $k \in \overset{o}{S}$. We have already proved in (P - 4.5) the existence (and the uniqueness comes by definition of the inverse operator) of

$$\left[Q_k^{-1}\right]^{-1} \quad \text{and} \quad \left[Q_k^{-1}(n)\right]^{-1}$$

for all $(k,n) \in \overset{o}{S} \times \bar{T}$. So we can define the event

$$C_k = \left\{ \lim_{n \to \infty} \left[Q_k^{-1}(n)\right]^{-1} \underline{q}_k(n) = \left[Q_k^{-1}\right]^{-1} \underline{q}_k \right\}$$

for each $k \in \overset{o}{S}$. Moreover,

$$A_k \cap B_k \subseteq C_k \;.$$

That is, the existence and uniqueness of $\left[Q_k^{-1}\right]^{-1}$ and $\left[Q_k^{-1}(n)\right]^{-1}$

for all $(k,n) \in \overset{o}{S} \times \bar{T}$, ensure that the occurence of $A_k \cap B_k$ implies the occurence of C_k for each $k \in \overset{o}{S}$. Hence

$$P(C_k) \geq P(A_k \cap B_k) = 1$$

$$\therefore \quad P(C_k) = P\{\lim_{n \to \infty} [Q_k^{-1}(n)]^{-1} \underline{g}_k(n) = [Q_k^{-1}]^{-1} \underline{g}_k\} = 1$$

for each $k \in \overset{o}{S}$. But

$$\underline{a}_k = Q_k \underline{g}_k = [Q_k^{-1}]^{-1} \underline{g}_k \qquad \text{(by lemma (L - 4.1))}$$

$$\underline{a}_k(n) = Q_k(n) \underline{g}_k(n) = [Q_k^{-1}(n)]^{-1} \underline{g}_k(n) \qquad \text{(by definition)}$$

for all $(k,n) \in \overset{o}{S} \times \bar{T}$. Then:

$$P(C_k) = P\{\lim_{n \to \infty} \underline{a}_k(n) = \underline{a}_k\} = 1$$

for each $k \in \overset{o}{S}$. $\square$

<u>Remarks</u>:

1) (Constant parameters) Consider the particular case where the parameters $\{\alpha_m \,;\; m = 1,\dots,M\}$ in (1) are constant over $X = (0,\ell)$. By (9) we can see that the components of $\underline{a}_k$, the parameter vector defined in (P - 4.1), will be constant over all $k \in S$ (i.e., $\underline{a}_k = \underline{a} \in R^{M+1}$ for all $k \in S$: a space-invariant vector in R^{M+1}). Thus, in order to perform the identification of $\underline{a}$ using the algorithm (SA-1), it will be sufficient to take noisy measurements $\underline{z}_{k_o}(n+N_t)$ in just M+1 points, say $\{x_{k_o-\tilde{M}},\dots,x_{k_o},\dots,x_{k_o+\bar{M}}\}$, located in the spatial domain X; where k_o is any fixed point in $\overset{o}{S}$. Moreover, the condition

$$\beta(x) \neq 0 \,; \qquad \forall\, x \in X$$

can be replaced by

$$\beta_k = \beta(x_k) \neq 0 \ ; \qquad \forall \ k = k_o - \tilde{M}, \ldots, k_o, \ldots, k_o + \bar{M}$$

since, in this case, it is just required that the particular matrix $Q_{k_o} \in BL(R^{M+1})$ is positive definite (see (P - 4.3) and its proof).

2) (No noise condition) If d=0 in the observation equation (2) we get $\underline{\varepsilon}_k = \underline{0}$, and so the identification algorithm (SA-1) does not depend on the knowledge of $\{\delta_m(x_k) \ ; \ m = 0, \ldots, N\}$.

3) (An indispensable information) Finally we recall that the information given in (11), that is:

$$\sum_{m=1}^{M+1} a_k^m \quad \text{independent of } \{\alpha_m(x_k) \ ; \ m = 1, \ldots, M\} \text{ for any } k \in S,$$

represented a fundamental step in our identification procedure (when $d \neq 0$), since it ensured that $\nu_k(\tau_n)$ in (P - 4.1) does not depend on $\underline{a}_k$, and so the result of lemma (L - 4.1) could be achieved.

4.5 - RECOVERING THE ORIGINAL PARAMETERS

In this section we consider the 2nd step of the identification procedure. That is, we face the problem of determining the set of parameters $\{\alpha_m(x_k) \ ; \ m = 1, \ldots, M\}$ appearing in the distributed model (1-b), for each $k \in \overset{o}{S}$.

To begin with, we recall the equation (9):

$$a_k^{i+\tilde{M}+1} = \sum_{m=m_\alpha(i)}^{M} (-1)^{\bar{m}-i} \binom{m}{i+\tilde{m}} \frac{\alpha_m(x_k)}{\delta_x^m} \qquad (9)$$

for each $i = -\tilde{M}, -M+1, \ldots, \bar{M}$ and $k \in S$, where

$$m_\alpha(i) = \begin{cases} -2i & ; \quad \text{if } i \le -1 \\ 1 & ; \quad \text{if } i = 0 \\ 2i-1 & ; \quad \text{if } i \ge 1 \end{cases}$$

Or equivalently:

$$a_k^{i+\tilde{M}+1} = \sum_{m=1}^{M} r_{im} \, \alpha_m(x_k) \tag{9'}$$

where:

$$r_{im} = \begin{cases} (-1)^{\bar{m}-i} \binom{m}{i+\tilde{m}} \dfrac{1}{\delta_x^m} & ; \quad \text{if } m \ge m_\alpha(i) \\ 0 & ; \quad \text{if } m < m_\alpha(i) \end{cases}$$

Now define

$$\underline{\alpha}_k = (\alpha_1(x_k), \ldots, \alpha_M(x_k)) \in R^M \; ; \qquad k \in S$$

$$R = \left[r_{im} \right] : \; R^M \to R^{M+1}$$

and recall

$$\underline{a}_k = (a_k^1, \ldots, a_k^{M+1}) \in R^{M+1} \; ; \qquad k \in S \, .$$

Then, by (9'), we get

$$\underline{a}_k = R \, \underline{\alpha}_k \; ; \qquad k \in S \; . \tag{18}$$

The problem here is to recover the original parameters $\underline{\alpha}_k \in R^M$ from $\underline{a}_k \in R^{M+1}$, for each $k \in \mathring{S}$. Since the stochastic approximation algorithm (SA-1) give us an estimate $\underline{a}_k(n)$ of $\underline{a}_k$ for all $n \in \bar{T}$ and for each $k \in \mathring{S}$, an estimate $\underline{\alpha}_k(n)$ of $\underline{\alpha}_k$ is then supplied by means of equation (18) for all $n \in \bar{T}$ and for each $k \in \overset{\circ}{S}$:

$$\underline{a}_k(n) = R\,\underline{\alpha}_k(n) \;; \qquad k \in \overset{\circ}{S} , \qquad n \in \bar{T} \;. \tag{19}$$

But $\underline{a}_k(n)$ and $\underline{\alpha}_k(n)$ take values in R^{M+1} and R^M, respectively. So it is not always possible to obtain a vector $\underline{\alpha}_k(n)$ exactly satisfying the equation (19). A standard alternative consist of determining an estimate $\hat{\underline{\alpha}}_k(n)$ of $\underline{\alpha}_k(n)$ which best approximates a solution in the sense of minimizing the norm $\| \underline{a}_k(n) - R\,\underline{\alpha}_k(n) \|$ over all $\underline{\alpha}_k(n)$. That is: a simple least-squares approach.

<u>Lemma (L - 4.3)</u>: (Least-Squares Estimate) The estimate of $\underline{\alpha}_k(n)$ which minimizes $\| \underline{a}_k(n) - R\,\underline{\alpha}_k(n) \|$ for each $(k,n) \in \overset{\circ}{S} \times \bar{T}$ is given by

$$\hat{\underline{\alpha}}_k(n) = (R^* R)^{-1} R^* \underline{a}_k(n) \;.$$

<u>Proof</u>:

By a direct inspection on the matrix R we can conclude that it has linearly independent columns, and so the proof comes as in [11] pp. 83. □

<u>Theorem (T - 4.2)</u>: The random sequence in $R^M \{ \hat{\underline{\alpha}}_k(n) \;;\; n \in \bar{T} \}$ obtained in (L - 4.3) converges to

$$\underline{\alpha}_k = (\alpha_1(x_k), \ldots, \alpha_M(x_k)) \qquad \in \; R^M$$

with probability one, for each $k \in \overset{\circ}{S}$:

$$P\{ \lim_{n \to \infty} \hat{\underline{\alpha}}_k(n) = \underline{\alpha}_k \} = 1 \;; \qquad k \in \overset{\circ}{S} \;.$$

Proof:

By (T - 4.1), (L - 4.3) and (18) we get :

$$\hat{\underline{\alpha}}_k(n) = (R^* R)^{-1} R^* \underline{a}_k(n) \xrightarrow[n \to \infty]{w.p.1} (R^* R)^{-1} R^* \underline{a}_k = (R^* R)^{-1} R^* R \underline{\alpha}_k = \underline{\alpha}_k . \square$$

4.6 - AN EQUIVALENT PROCEDURE

We present here an equivalent procedure for the identification method developed in sections 4.2 through 4.5. It is based on formulating a slightly different version of the proposition (P - 4.1), as follows: If we define

i) $$z_k(\tau_n') = \frac{1}{c_k^{\tilde{N}+1}} \left(z_k(\tau_n) - c_k^{\tilde{N}+1} z_k(n)\right) =$$

$$= \sum_{\substack{m=1 \\ m \neq \tilde{N}+1}}^{N} \frac{c_k^m}{c_k^{\tilde{N}+1}} z_k(n+m-\tilde{N}-1) + z_k(n+\tilde{N})$$

ii) $$\nu_k(\tau_n') = \frac{1}{c_k^{\tilde{N}+1}} \nu_k(\tau_n)$$

iii) $$\underline{a}_k' = \frac{1}{c_k^{\tilde{N}+1}} \left(\underline{a}_k - c_k^{\tilde{N}+1} \underline{e}_{\tilde{M}+1}\right) =$$

$$= \frac{1}{c_k^{\tilde{N}+1}} \left(a_k^1, \ldots, a_k^{\tilde{M}}, a_k^{\tilde{M}+1} - c_k^{\tilde{N}+1}, a_k^{\tilde{M}+2}, \ldots, a_k^{M+1}\right) \quad \in \ R^{M+1}$$

where:

$$\underline{e}_{\tilde{M}+1} = (0, \ldots, 0, 1, 0, \ldots, 0) \quad \in \ R^{M+1}$$

with 1 in the $(\tilde{M}+1)$th position and zeros elsewhere.

iv) $\underline{\varepsilon}_k' = \frac{1}{c_k^{N+1}} \underline{\varepsilon}_k \qquad \in \quad R^{M+1}$

v) $\underline{q}_k' = E\{\underline{z}_k(n+N_t)\ z_k(\tau'_{n+N_t})\} + \underline{\varepsilon}_k' \qquad \in \quad R^{M+1}$

the results in (P - 4.1) and (L - 4.1) take the following form:

(P - 4.1): $\quad z_k(\tau_n') = \langle \underline{a}_k' \,;\, \underline{z}_k(n) \rangle + \nu_k(\tau_n')$,

(L - 4.1): $\quad \underline{a}_k' = Q_k\, \underline{q}_k'$,

and a similar version of the identification algorithm presented in (T - 4.1) is obtained when τ_n, $\underline{\varepsilon}_k$, and $\underline{a}_k$ are replaced by τ_n', $\underline{\varepsilon}_k'$ and $\underline{a}_k'$, respectively:

(T - 4.1): $\quad \underline{a}_k'(n) \xrightarrow[n \to \infty]{w.p.1} \underline{a}_k'$

where:

$$\underline{a}_k'(n) = \underline{a}_k(n) \Big|_{\substack{\tau_n = \tau_n' \\ \underline{\varepsilon}_k = \underline{\varepsilon}_k'}} \qquad \text{in (SA-1)} .$$

Now define:

vi) $\underline{\alpha}_k' = \frac{1}{c_k^{N+1}} \left(\frac{\alpha_1(x_k)}{\delta_x}, \ldots, \frac{\alpha_M(x_k)}{\delta_x^M} \right) \qquad \in \quad R^M$

vii) $r_{im}' = r_{im}\, \delta_x^m$

viii) $R' = \left[r_{im}' \right] \ : \ R^M \rightarrow R^{M+1}$

So we get, from (9'):

$$\underline{a}_k = R\, \underline{\alpha}_k = c_k^{N+1}\, R'\, \underline{\alpha}_k'$$

and since

$$\underline{a}'_k = \frac{\underline{a}_k}{c_k^{N+1}} - \frac{c_k^{\tilde{N}+1}}{c_k^{N+1}} \underline{e}_{\tilde{M}+1}$$

the equations (18) and (19) become:

$$(18): \quad \underline{a}'_k = R' \underline{\alpha}'_k - \frac{c_k^{\tilde{N}+1}}{c_k^{N+1}} \underline{e}_{\tilde{M}+1}$$

$$(19): \quad \underline{a}'_k(n) = R' \underline{\alpha}'_k(n) - \frac{c_k^{\tilde{N}+1}}{c_k^{N+1}} \underline{e}_{\tilde{M}+1}$$

In this way the least-squares estimate of $\underline{\alpha}'_k(n)$ is given by:

$$(L - 4.3): \quad \hat{\underline{\alpha}}'_k(n) = (R'^{*} R')^{-1} R'^{*} \left[\underline{a}'_k(n) + \frac{c_k^{\tilde{N}+1}}{c_k^{N+1}} \underline{e}_{\tilde{M}+1} \right] ,$$

and the convergence is proved as in (T - 4.2) by using (18), (T - 4.1) and (L - 4.3):

$$(T - 4.2): \quad \hat{\underline{\alpha}}'_k(n) \xrightarrow[n \to \infty]{w.p.1} \underline{\alpha}'_k \; .$$

This equivalent procedure was applied in [4] for a particular classof second-order models. Note that the original procedure identifies parameters appearing in the discrete version given by (5) (i.e., parameters of the matrix A), while the equivalent procedure identifies unknown parameters appearing in the system matrix F given in (13) (more precisely, parameters of the matrix $A_{\tilde{N}+1} = C_{N+1}^{-1} (A - C_{\tilde{N}+1}))$.

and since

$$\underline{a}'_k = \frac{\underline{a}_k}{c_k^{N+1}} - \frac{c_k^{\tilde{N}+1}}{c_k^{N+1}}\,\underline{e}_{\tilde{M}+1}$$

the equations (18) and (19) become:

$$(18): \qquad \underline{a}'_k = R'\,\underline{\alpha}'_k - \frac{c_k^{\tilde{N}+1}}{c_k^{N+1}}\,\underline{e}_{\tilde{M}+1}$$

$$(19): \qquad \underline{a}'_k(n) = R'\,\underline{\alpha}'_k(n) - \frac{c_k^{\tilde{N}+1}}{c_k^{N+1}}\,\underline{e}_{\tilde{M}+1}$$

In this way the least-squares estimate of $\underline{\alpha}'_k(n)$ is given by:

$$(L - 4.3): \qquad \hat{\underline{\alpha}}'_k(n) = (R'^{*}R')^{-1}\,R'^{*}\left[\underline{a}'_k(n) + \frac{c_k^{\tilde{N}+1}}{c_k^{N+1}}\,\underline{e}_{\tilde{M}+1}\right],$$

and the convergence is proved as in (T - 4.2) by using (18), (T - 4.1) and (L - 4.3):

$$(T - 4.2): \qquad \hat{\underline{\alpha}}'_k(n) \xrightarrow[n \to \infty]{w.p.1} \underline{\alpha}'_k \ .$$

This equivalent procedure was applied in [4] for a particular classof second-order models. Note that the original procedure identifies parameters appearing in the discrete version given by (5) (i.e., parameters of the matrix A), while the equivalent procedure identifies unknown parameters appearing in the system matrix F given in (13) (more precisely, parameters of the matrix $A_{\tilde{N}+1} = C_{N+1}^{-1}\,(A - C_{\tilde{N}+1})$).

Remark: (No noise condition) We can show that, if d=0 in (2), the stochastic approximation algorithm in (T - 4.1) also identifies

$$c_k^{\tilde{N}+1} = \sum_{m=0}^{N} (-1)^{\bar{m}-j} \binom{m}{j+\tilde{m}} \frac{\gamma_m(x_k)}{\delta_t^m} .$$

Set

$$R'' = \left[R' \;\vdots\; -\underline{e}_{\tilde{M}+1} \right] \quad \in \; BL(R^{M+1})$$

and note that there exists R''^{-1}, since the independent columns of R' are also independent of $\underline{e}_{\tilde{M}+1}$. Now defining

$$\underline{\alpha}_k'' = \left(\underline{\alpha}_k', \frac{c_k^{\tilde{N}+1}}{c_k^{N+1}}\right) = \frac{1}{c_k^{N+1}} \left(\frac{\alpha_1(x_k)}{\delta_x}, \ldots, \frac{\alpha_M(x_k)}{\delta_x^M}, c_k^{\tilde{N}+1}\right) \quad \in \; R^{M+1}$$

the equation (18) becomes:

$$\underline{a}_k' = R'' \underline{\alpha}_k'' \qquad \in \; R^{M+1}$$

and so we get a recursive estimate of $\underline{\alpha}_k''$ directly from (T - 4.1)

$$\underline{\alpha}_k''(n) = R''^{-1} \underline{a}_k'(n) ,$$

since $\underline{a}_k'(n)$ does not depend on $\{\gamma_m(x_k) \; ; \; m = 0, \ldots, N\}$ if d=0.

4.7 - EXTENSION TO MULTI-DIMENSIONAL SPATIAL DOMAIN

So far we have been considering a one-dimensional spatial domain ($x \in X \subset R^1$). Direct extensions of the theory developed in the previous sections can be obtained for distributed models involving multi-dimensional spatial domains ($x \in X \subset R^p$). In order to illustrate this, we perform the main steps of the identification procedure for the following second-order (M=2, N=1) linear model with two independent spatial variables ($x \in X \subset R^2$) and constant coefficients:

$$\gamma_0\, u(x_1,x_2,t) + \gamma_1 \frac{\partial}{\partial t} u(x_1,x_2,t) =$$

$$= \alpha_{11} \frac{\partial}{\partial x_1} u(x_1,x_2,t) + \alpha_{12} \frac{\partial}{\partial x_2} u(x_1,x_2,t) +$$

$$+ \alpha_{21} \frac{\partial^2}{\partial x_1^2} u(x_1,x_2,t) + \alpha_{22} \frac{\partial^2}{\partial x_2^2} u(x_1,x_2,t) +$$

$$+ \beta(x_1,x_2)\, w_{x_1,x_2}(t)\,.$$

Since the solution method follows exactly as before, we will omit any comments and just the basic results will be presented in a concise form. The notation remains the same as in the preceding sections and, for simplicity, we consider homogeneous boundary conditions. The spatial domain X is taken to be an open square in R^2.

<u>Continuous formulation</u>: Infinite-dimensional state space.

MODEL: $$\sum_{m=0}^{1} \gamma_m \frac{\partial^m}{\partial t^m} u(x,t) = \sum_{m=1}^{2} \sum_{i=1}^{2} \alpha_{mi} \frac{\partial^m}{\partial x_i^m} u(x,t) + \beta(x)\, w_x(t)$$

$$x = (x_1,x_2) \in X = (0,\ell) \times (0,\ell) \subset R^2 \;; \qquad t > 0$$

IC: $u(x,0) = g(x) \;; \qquad x \in \bar{X} = [0,\ell] \times [0,\ell]$

BC: $u(x',t) = 0 \;; \qquad x' \in \Gamma$: the boundary of X ; $\qquad t \geq 0$

OBS: $z(x_k,t) = h\, u(x_k,t) + d\, v(t)$

$x_k \in P_x$: a partition of $\bar{X} \subset R^2$; $\qquad t > 0$

where:

i) The input disturbances $\{w_x(t)\ ;\ t \geq 0\}$ are taken to be real-valued second-order stochastic processes for each x in $(0,\ell)\times(0,\ell)\subset R^2$.

ii) The observation noise $\{v(t)\ ;\ t > 0\}$ is a stochastic process defined as before.

iii) $\{\alpha_{mi}\ ;\ m,i=1,2\}$, $\{\gamma_m\ ;\ m=0,1\}$, h and d are assumed to be real constant parameters; and $\gamma_1 \neq 0$, $h \neq 0$.

iv) $\beta(x) \neq 0$; for all $x \in X$.

<u>Space-time discretization</u>: Finite-dimensional discrete version.

MODEL:
$$c_1 u_{k_1,k_2}(n) + c_2 u_{k_1,k_2}(n+1) = a_1 u_{k_1-1,k_2}(n) + a_{21} u_{k_1,k_2-1}(n) + a_{22} u_{k_1,k_2}(n) + a_{23} u_{k_1,k_2+1}(n) + a_3 u_{k_1+1,k_2}(n) + \beta_{k_1,k_2} w_{k_1,k_2}(n)$$
$$(k_1,k_2)\in S\times S\ ; \quad n\in \bar{T}$$

IC:
$$u_{k_1,k_2}(0) = g(x_{k_1},x_{k_2})\ ; \quad (k_1,k_2)\in S\times S$$

BC:
$$\left.\begin{array}{l} u_{0,k_2}(n) = u_{K+1,k_2}(n) = 0 \\ u_{k_1,0}(n) = u_{k_1,K+1}(n) = 0 \end{array}\right\} ; \quad (k_1,k_2)\in \bar{S}\times\bar{S}\ ; \quad n\in\bar{T}$$

OBS:
$$z_{k_1,k_2}(n) = h\, u_{k_1,k_2}(n) + d\, v(n)$$
$$(k_1,k_2)\in S\times S\ ; \quad n\in T$$

where:

i) $x_{k_i} = k_i \, \delta_{x_i} \in (0,\ell)$; $i = 1,2$ (see fig. 7)

i-a) $\delta_{x_1} = \delta_{x_2} = \delta_x$

i-b) $k_i \in \bar{S} = \{0,1,\ldots,K+1\} \supset S = \{1,2,\ldots,K\}$

i-c) $K = \dfrac{\ell}{\delta_x} - 1$

i-d) $\delta_x > 0$ such that K is an integer ≥ 3

ii) $t_n = n \, \delta_t \geq 0$

ii-a) $n \in \bar{T} = \{0,1,\ldots\} \supset T = \{1,2,\ldots\}$

ii-b) $\delta_t > 0$

iii) $u_{k_1,k_2}(n) = u(x_{k_1}, x_{k_2}, t_n)$

iii-a) $\dfrac{\partial}{\partial t} u(x_{k_1}, x_{k_2}, t)\Big|_{t=t_n} \simeq D_{\delta_t} u(x_{k_1}, x_{k_2}, t_n) =$

$$= \frac{1}{\delta_t} \left[u_{k_1,k_2}(n+1) - u_{k_1,k_2}(n) \right]$$

iii-b) $\dfrac{\partial}{\partial x_i} u(x,t_n)\Big|_{x_i = x_{k_i}} \simeq D_{\delta_{x_i}} u(x_{k_1}, x_{k_2}, t_n) =$

$$= \frac{1}{\delta_x} \begin{cases} \left[u_{k_1+1,k_2}(n) - u_{k_1,k_2}(n) \right] & ; \quad i = 1 \\ \left[u_{k_1,k_2+1}(n) - u_{k_1,k_2}(n) \right] & ; \quad i = 2 \end{cases}$$

$$\text{iii-c)}\quad \frac{\partial^2}{\partial x_i^2} u(x,t_n)\Big|_{x_i=x_{k_i}} \cong D^2_{\delta_{x_i}} u(x_{k_1},x_{k_2},t_n) =$$

$$= \frac{1}{\delta_x^2}\begin{cases}\left[u_{k_1-1,k_2}(n) - 2\,u_{k_1,k_2}(n) + u_{k_1+1,k_2}(n)\right] \;; & i=1\\[2ex] \left[u_{k_1,k_2-1}(n) - 2\,u_{k_1,k_2}(n) + u_{k_1,k_2+1}(n)\right] \;; & i=2\end{cases}$$

$$\text{iv)}\quad w_{k_1,k_2}(n) = w_{x_{k_1},x_{k_2}}(t_n)$$

$$\text{v)}\quad v(n) = v(t_n)$$

$$\text{vi)}\quad z_{k_1,k_2}(n) = z(x_{k_1},x_{k_2},t_n)$$

$$\text{vii)}\begin{cases} a_1 = \dfrac{1}{\delta_x^2}\alpha_{21}\\[2ex] a_{21} = \dfrac{1}{\delta_x^2}\alpha_{22}\\[2ex] a_{22} = -\dfrac{1}{\delta_x}\left(\alpha_{11} + \alpha_{12} + \dfrac{2\alpha_{21}}{\delta_x} + \dfrac{2\alpha_{22}}{\delta_x}\right)\\[2ex] a_{23} = \dfrac{1}{\delta_x}\left(\alpha_{12} + \dfrac{\alpha_{22}}{\delta_x}\right)\\[2ex] a_3 = \dfrac{1}{\delta_x}\left(\alpha_{11} + \dfrac{\alpha_{21}}{\delta_x}\right)\end{cases}$$

$$\text{viii)}\begin{cases} c_1 = \gamma_0 - \dfrac{\gamma_1}{\delta_t}\\[2ex] c_2 = \dfrac{\gamma_1}{\delta_t}\end{cases}$$

$$\text{ix)}\quad \beta_{k_1,k_2} = \beta(k_1,k_2)$$

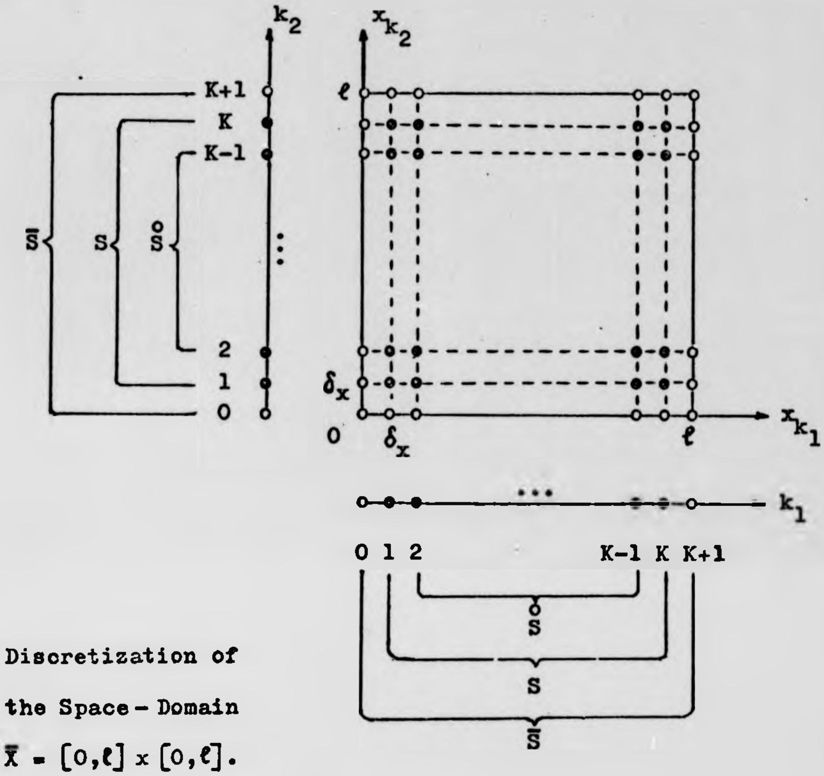

Fig. 7: Discretization of the Space-Domain $\bar{X} = [0,\ell] \times [0,\ell]$.

Equivalent discrete-time LPS:

$$\begin{cases} \text{MODEL:} & \underline{u}(n+1) = A\,\underline{u}(n) + B\,\underline{w}(n)\ ; \qquad \underline{u}(0) \text{ given} \\ \text{OBS:} & \underline{z}(n) = h\,\underline{u}(n) + \underline{d}\,v(n)\ ; \qquad n \in T \end{cases}$$

where:

i) $\underline{u}(n) = (u_{11}(n),\dots,u_{1K}(n),\dots\dots,u_{K1}(n),\dots,u_{KK}(n))\ ; \qquad n \in \bar{T}$

ii) $\underline{w}(n) = (w_{11}(n),\dots,w_{1K}(n),\dots\dots,w_{K1}(n),\dots,w_{KK}(n))\ ; \qquad n \in \bar{T}$

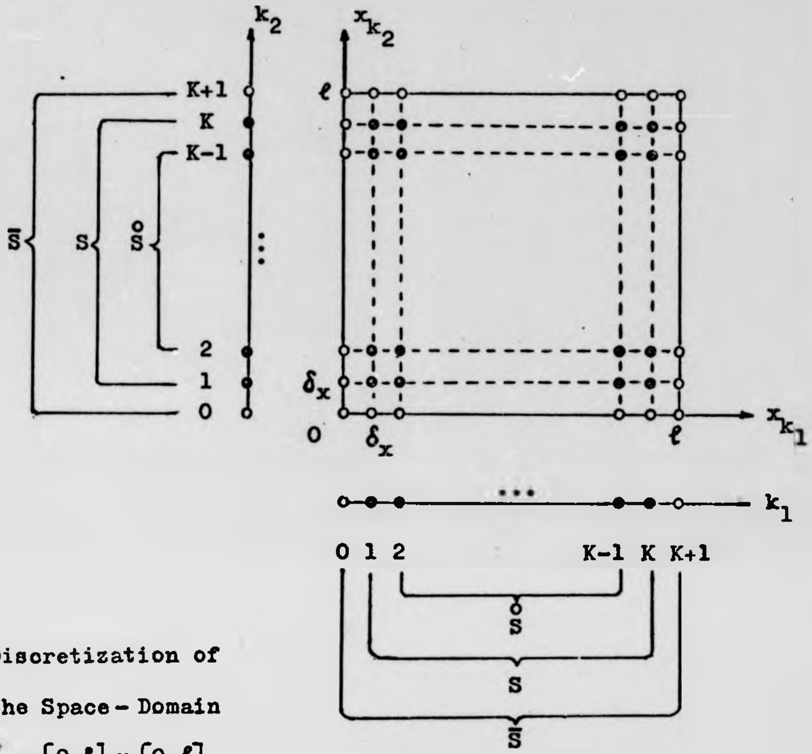

Fig. 7: Discretization of the Space – Domain $\bar{X} = [0,\ell] \times [0,\ell]$.

Equivalent discrete-time LPS:

$$\begin{cases} \text{MODEL:} & \underline{u}(n+1) = A\,\underline{u}(n) + B\,\underline{w}(n)\,; & \underline{u}(0) \text{ given} \\ \text{OBS:} & \underline{z}(n) = h\,\underline{u}(n) + \underline{d}\,v(n)\,; & n \in T \end{cases}$$

where:

i) $\underline{u}(n) = (u_{11}(n),\ldots,u_{1K}(n),\ldots\ldots,u_{K1}(n),\ldots,u_{KK}(n))\,; \qquad n \in \bar{T}$

ii) $\underline{w}(n) = (w_{11}(n),\ldots,w_{1K}(n),\ldots\ldots,w_{K1}(n),\ldots,w_{KK}(n))\,; \qquad n \in \bar{T}$

iii) $\underline{z}(n) = (z_{11}(n),\ldots,z_{1K}(n),\ldots\ldots,z_{K1}(n),\ldots,z_{KK}(n))\ ; \qquad n \in T$

random vectors in R^{K^2}.

iv) $\underline{d} = d(1,1,\ldots,1) \qquad \in \quad R^{K^2}$

v) $$A = \begin{bmatrix} A_2 & A_3 & & & \\ A_1 & A_2 & A_3 & & \\ & \ddots & \ddots & \ddots & \\ & & A_1 & A_2 & A_3 \\ & & & A_1 & A_2 \end{bmatrix} \in BL(R^{K^2})$$

vi) $A_1 = \dfrac{a_1}{c_2} I \qquad \in \ BL(R^K)$

vii) $A_3 = \dfrac{a_3}{c_2} I \qquad \in \ BL(R^K)$

viii) $$A_2 = \frac{1}{c_2} \begin{bmatrix} a_{22} & a_{23} & & & \\ a_{21} & a_{22} & a_{23} & & \\ & \ddots & \ddots & \ddots & \\ & & a_{21} & a_{22} & a_{23} \\ & & & a_{21} & a_{22} \end{bmatrix} - \frac{c_1}{c_2} I \qquad \in \ BL(R^K)$$

ix) $$B = \frac{1}{c_2} \begin{bmatrix} \beta_{11} & & & & \\ & \ddots & & & \\ & & \beta_{1K} & & \\ & & & \ddots & \\ & & & & \beta_{K1} & \\ & & & & & \ddots & \\ & & & & & & \beta_{KK} \end{bmatrix} \in \ BL(R^{K^2})$$

Identification Procedure: Since in this case we still have

$$a_1 + \sum_{i=1}^{3} a_{2i} + a_3 = 0 ,$$

a relation expressing the observation dynamics, as presented in (P - 4.1), can be obtained as follows:

Define

i) $\overset{\circ}{S} = \{2,\ldots,K-1\} \subset S$

ii) $z_{k_1,k_2}(\tau_n) = c_1\, z_{k_1,k_2}(n) + c_2\, z_{k_1,k_2}(n+1)$

iii) $\nu_{k_1,k_2}(\tau_n) = d\,(c_1\, v(n) + c_2\, v(n+1)) + h\, \beta_{k_1,k_2}\, w_{k_1,k_2}(n)$

iv) $\underline{a} = (a_1, a_{21}, a_{22}, a_{23}, a_3) \quad \in \quad R^5$

v) $\underline{z}_{k_1,k_2}(n) = (z_{k_1-1,k_2}(n), z_{k_1,k_2-1}(n), z_{k_1,k_2}(n), z_{k_1,k_2+1}(n),$

$z_{k_1+1,k_2}(n))$: randon vectors in R^5.

Proposition (P - 4.1'): For each $(k_1,k_2) \in \overset{\circ}{S} \times \overset{\circ}{S}$ and $n \in T$,

$$z_{k_1,k_2}(\tau_n) = \langle\, \underline{a}\ ,\ \underline{z}_{k_1,k_2}(n)\, \rangle + \nu_{k_1,k_2}(\tau_n) .$$

Based on (P - 4.1') it can be shown that the parameter explicit lemma (L - 4.1) has a similar version in case of multi-dimensional spatial domain. First consider the assumptions (A - 4.1) - (A - 4.4) (stability, stationarity, steady state and finite transient time), where the input disturbance $\{\underline{w}(n)\ ;\ n \in \overline{T}\}$ is a random sequence in R^{K^2} and so C_w is a symmetric positive definite matrix in $BL(R^{K^2})$. Now define

i) $\varepsilon = -d^2 \left[\eta_v^2 \, \gamma_0 + c_1 \, \sigma_v(0) + c_2 \, \sigma_v(1)\right]$

ii) $\underline{\varepsilon} = \varepsilon(1,1,\ldots,1) \quad \in \; R^5$

iii) $\underline{q}_{k_1,k_2} = E\{z_{k_1,k_2}(\zeta_{n+N_t}) \; \underline{z}_{k_1,k_2}(n+N_t)\} + \underline{\varepsilon} \quad \in \; R^5$

iv) $Q^{-1}_{k_1,k_2} = E\{\underline{z}_{k_1,k_2}(n+N_t) \; \underline{z}^{*}_{k_1,k_2}(n+N_t)\} \quad \in \; BL(R^5)$

$$\text{v)} \quad J_{k_1,k_2} = \left[\begin{array}{ccc:c:ccc} 0\ldots0 & 1 & 0\ldots & & & & \ldots 0 \\ \hdashline & \emptyset & & I & & \emptyset & \\ \hdashline 0\ldots & & & & \ldots 0 & 1 & 0\ldots 0 \end{array}\right] \; : \; R^{K^2} \longrightarrow R^5$$

with $I \in BL(R^3)$ centered at the $[(k_1-1)K+k_2]$ th position and the 1s, in the first and fifth rows, placed at $[(k_1-2)K+k_2]$ th and $[k_1 K+k_2]$ th positions, respectively.

Since

$$\underline{z}_{k_1,k_2}(n) = J_{k_1,k_2} \, \underline{z}(n)$$

the propositions (P - 4.2) - (P - 4.5) remain valid if we replace $k \in \overset{\circ}{S}$ by $(k_1,k_2) \in \overset{\circ}{S} \times \overset{\circ}{S}$. Hence:

<u>Lemma (L - 4.1')</u>: (Explicit parameter) Let $\underline{q}_{k_1,k_2}$ and Q_{k_1,k_2} be as defined before. If the assumptions (A - 4.1) - (A - 4.4) are satisfied, then the parameter vector

$$\underline{a} = (a_1, a_{21}, a_{22}, a_{23}, a_3) \quad \in \; R^5$$

introduced in (P - 4.1') can be placed in an explicit form, as follows:

$$\underline{a} = Q_{k_1,k_2} \, \underline{q}_{k_1,k_2}$$

i) $\varepsilon = -d^2 \left[\eta_v^2 \, \gamma_0 + c_1 \, \delta_v(0) + c_2 \, \delta_v(1)\right]$

ii) $\underline{\varepsilon} = \varepsilon(1,1,\ldots,1) \quad \in \; R^5$

iii) $\underline{q}_{k_1,k_2} = E\{z_{k_1,k_2}(z_{n+N_t}) \, \underline{z}_{k_1,k_2}(n+N_t)\} + \underline{\varepsilon} \quad \in \; R^5$

iv) $Q^{-1}_{k_1,k_2} = E\{\underline{z}_{k_1,k_2}(n+N_t) \, \underline{z}^*_{k_1,k_2}(n+N_t)\} \qquad \in \; BL(R^5)$

v) $$J_{k_1,k_2} = \begin{bmatrix} 0\ldots0\;1\;0\ldots & & \ldots0 \\ \hline \emptyset & I & \emptyset \\ \hline 0\ldots & & \ldots0\;1\;0\ldots0 \end{bmatrix} \; : \; R^{K^2} \longrightarrow R^5$$

with $I \in BL(R^3)$ centered at the $[(k_1-1)K+k_2]$ th position and the 1s, in the first and fifth rows, placed at $[(k_1-2)K+k_2]$ th and $[k_1 K+k_2]$ th positions, respectively.

Since

$$\underline{z}_{k_1,k_2}(n) = J_{k_1,k_2} \, \underline{z}(n)$$

the propositions (P - 4.2) - (P - 4.5) remain valid if we replace $k \in \overset{o}{S}$ by $(k_1,k_2) \in \overset{o}{S} \times \overset{o}{S}$. Hence:

<u>Lemma (L - 4.1')</u>: (Explicit parameter) Let $\underline{q}_{k_1,k_2}$ and Q_{k_1,k_2} be as defined before. If the assumptions (A - 4.1) - (A - 4.4) are satisfied, then the parameter vector

$$\underline{a} = (a_1, a_{21}, a_{22}, a_{23}, a_3) \quad \in \; R^5$$

introduced in (P - 4.1') can be placed in an explicit form, as follows:

$$\underline{a} = Q_{k_1,k_2} \, \underline{q}_{k_1,k_2}$$

for any $(k_1,k_2) \in \overset{o}{S} \times \overset{o}{S}$.

In this way, the parameter vector $\underline{a}$ can be identified through noisy observations $z_{k_1,k_2}(n)$, by using stochastic approximation algorithms as presented in (T - 4.1) (with $k \in \overset{o}{S}$ replaced by $(k_1,k_2) \in \overset{o}{S} \times \overset{o}{S}$).

Remarks:

1) As far as the equivalent LPS is concerned, the computational complexity increases exponentially with the dimension of the spatial domain X. For instance, assume X is an open rectangle in R^p, and let K be a fixed integer ($\geq M+1$) such that the discretization of X contains K^p interior points (i.e, K interior points for each discretized coordinate: $S = \{1,2,\ldots,K\}$). So, as shown for the two-dimensional case (with $N = 1$), the equivalent LPS will be of order $(NK)^p$ (i.e., $\underline{y}(n)$ is in $R^{(NxK)^p}$).

2) For identification purposes, the computational complexity increases only with the number of parameters to be identified (i.e., the identification algorithm $\underline{a}_k(n)$ is in $R^{p^{M+1}}$ for each k).

CHAPTER 5

SUMMARY, EXAMPLES AND CONSIDERATIONS

The performance of the identification method proposed in chapter 4 is analysed. After a brief summary concerning second-order models, we present three examples dealing with parabolic and hyperbolic PDE. The chapter closes with a concise list of remarks including some conclusions and suggestions for further research in this field.

5.1 - SUMMARY: SECOND-ORDER MODELS

We present here a brief summary of the identification procedure developed in the last chapter, for second-order models ($M = 2$, $0 < N \leq 2$) with constant (space-invariant) parameters and one-dimensional spatial domain ($x \in (0,\ell)$). Two cases will be considered separately: $M = 2$, $N = 1$ and $M = N = 2$.[1]

---

[1] Recall that

$$M = 2 \longrightarrow \tilde{M} = \bar{M} = 1$$

$$N = \begin{cases} 1 & \longrightarrow \tilde{N} = 0 \ , \quad \bar{N} = 1 \\ 2 & \longrightarrow \tilde{N} = \bar{N} = 1 \end{cases}$$

Original DPS:

$$
\begin{cases}
\text{MODEL:} & \left.\begin{array}{ll} \underline{N=1}: & \gamma_0\, u(x,t) + \gamma_1 \dfrac{\partial}{\partial t} u(x,t) \\ \underline{N=2}: & \gamma_0\, u(x,t) + \gamma_1 \dfrac{\partial}{\partial t} u(x,t) + \gamma_2 \dfrac{\partial^2}{\partial t^2} u(x,t) \end{array}\right\} = \\
& = \alpha_1 \dfrac{\partial}{\partial x} u(x,t) + \alpha_2 \dfrac{\partial^2}{\partial x^2} u(x,t) + \beta(x)\, w_x(t) \ ; \quad x \in (0,\ell) \ ; \quad t > 0 \\
\text{IC:} & \begin{array}{ll} \underline{N=1}: & u(x,0) = g_0(x) \\ \underline{N=2}: & u(x,0) = g_0(x) \ ; \quad \dfrac{\partial}{\partial t} u(x,t)\Big|_{t=0} = g_1(x) \end{array} \ ; \quad x \in [0,\ell] \\
\text{BC:} & u(0,t) = u(\ell,t) = 0 \ ; \quad t \geq 0 \\
\text{OBS:} & z(x_k,t) = h\, u(x,t) + d\, v(t) \ ; \quad x_k \in P_x \ ; \quad t > 0
\end{cases}
$$

Discrete version:

$$
\begin{cases}
\text{MODEL:} & \left.\begin{array}{ll} \underline{N=1}: & c_1\, u_k(n) + c_2\, u_k(n+1) \\ \underline{N=2}: & c_1\, u_k(n-1) + c_2\, u_k(n) + c_3\, u_k(n+1) \end{array}\right\} = \\
& = a_1\, u_{k-1}(n) + a_2\, u_k(n) + a_3\, u_k(n) + \beta_k\, w_k(n) \ ; \quad (k,n) \in S \times T' \\
\text{IC:} & \begin{array}{ll} \underline{N=1}: & u_k(0) = g_0(x_k) \\ \underline{N=2}: & u_k(0) = g_0(x_k) \ ; \quad u_1(1) = \delta_t\, g_1(x_k) + g_0(x_k) \end{array} \ ; \quad k \in S \\
\text{BC:} & u_0(n) = u_{K+1}(n) = 0 \ ; \quad n \in \overline{T} \\
\text{OBS:} & z_k(n) = h\, u_k(n) + d\, v(n) \ ; \quad (k,n) \in S \times T
\end{cases}
$$

where the coefficients $\{a_1, a_2, a_3\}$ and $\{c_1, c_2, c_3\}$ are given by:

$$a_1 = \frac{\alpha_2}{\delta_x^2}$$

$$a_2 = -\frac{\alpha_1}{\delta_x} - 2\frac{\alpha_2}{\delta_x^2}$$

$$a_3 = \frac{\alpha_1}{\delta_x} + \frac{\alpha_2}{\delta_x^2}$$

$$\underline{N = 1}: \begin{cases} c_1 = \gamma_0 - \dfrac{\gamma_1}{\delta_t} \\ c_2 = \dfrac{\gamma_1}{\delta_t} \end{cases}$$

$$\underline{N = 2}: \begin{cases} c_1 = \dfrac{\gamma_2}{\delta_t^2} \\ c_2 = \gamma_0 - \dfrac{\gamma_1}{\delta_t} - 2\dfrac{\gamma_2}{\delta_t^2} \\ c_3 = \dfrac{\gamma_1}{\delta_t} + \dfrac{\gamma_2}{\delta_t^2} \end{cases}$$

and the sets $\bar{T}$, T', T and S are defined as follows:

$$\bar{T} = \{0,1,2,\ldots\}$$

$$\underline{N = 1}: \quad T = \{1,2,\ldots\} \subset T' = \bar{T}$$

$$\underline{N = 2}: \quad T = \{2,3,\ldots\} \subset T' = \{1,2,\ldots\} \subset \bar{T}$$

$$S = \{1,2,\ldots,K\}$$

with

$$K = \frac{\ell}{\delta_x} - 1 : \quad \text{an integer} \geq 3.$$

Equivalent discrete-time LPS:

$$\begin{cases} \text{MODEL:} & \underline{y}(n+1) = F\,\underline{y}(n) + G\,\underline{w}(n)\ ; \quad \underline{y}(\tilde{N}) \text{ given} \\ \text{OBS:} & \underline{z}(n) = H\,\underline{y}(n) + \underline{d}\,v(n)\ ; \quad n \in T \end{cases}$$

where:

i) $\underline{u}(n) = (u_1(n),\dots,u_K(n))\ ; \quad n \in \bar{T}\ : \quad \text{in } R^K$

ii) $\underline{w}(n) = (w_1(n),\dots,w_K(n))\ ; \quad n \in T'\ : \quad \text{in } R^K$

iii) $\underline{z}(n) = (z_1(n),\dots,z_K(n))\ ; \quad n \in T\ : \quad \text{in } R^K$

iv) $\underline{d} = d(1,1,\dots,1) \quad \in R^K$

v) $$B = \begin{bmatrix} \beta_1 & & \\ & \ddots & \\ & & \beta_K \end{bmatrix} \in BL(R^K)$$

vi) $$A = \begin{bmatrix} a_2 & a_3 & & & \\ a_1 & a_2 & a_3 & & \\ & \ddots & \ddots & \ddots & \\ & & a_1 & a_2 & a_3 \\ & & & a_1 & a_2 \end{bmatrix} \in BL(R^K)$$

N = 1:

vii-a) $\underline{y}(n) = \underline{u}(n)\ ; \quad n \in T' = T\ : \quad \text{in } R^K$

$\underline{y}(\tilde{N}) = \underline{y}(0) = \underline{u}(0) = (g_0(x_1),\dots,g_0(x_K))$

viii-a) $F = A_1 = C_2^{-1}(C_1 - A) = \dfrac{1}{c_2}(A - c_1 I) =$

$$
= \begin{bmatrix} a'_2 & a'_3 & & & \\ a'_1 & a'_2 & a'_3 & & \\ & \ddots & \ddots & \ddots & \\ & & a'_1 & a'_2 & a'_3 \\ & & & a'_1 & a'_2 \end{bmatrix} \in BL(R^K)
$$

where:

$$a'_1 = \frac{a_1}{c_2} = \alpha'_2$$

$$a'_2 = \frac{a_2 - c_1}{c_2} = -\alpha'_1 - 2\alpha'_2 - \frac{c_1}{c_2} = 1 - (\alpha'_1 + 2\alpha'_2 + \delta_t \frac{\gamma_0}{\delta_1})$$

$$a'_3 = \frac{a_3}{c_2} = \alpha'_1 + \alpha'_2$$

with the coefficients α'_1 and α'_2 defined as follows:

$$\alpha'_1 = \frac{1}{c_2} \frac{\alpha_1}{\delta_x} = \frac{1}{\delta_1} \frac{\delta_t}{\delta_x} \alpha_1$$

$$\alpha'_2 = \frac{1}{c_2} \frac{\alpha_2}{\delta_x^2} = \frac{1}{\delta_1} \frac{\delta_t}{\delta_x^2} \alpha_2$$

ix-a) $G = B_1 = C_2^{-1} B = \frac{1}{c_2} B = \frac{\delta_t}{\delta_1} B \quad \in BL(R^K)$

x-a) $H = h\, I \quad \in BL(R^K)$

<u>N = 2</u>:

vii-b) $\underline{y}(n) = (\underline{u}(n-1), \underline{u}(n))$; $\quad n \in T'$: in R^{2K}

$\underline{y}(\tilde{N}) = \underline{y}(1) = (\underline{u}(0), \underline{u}(1)) =$

$= (g_0(x_1), \ldots, g_0(x_K), \delta_t\, g_1(x_1) + g_0(x_1), \ldots, \delta_t\, g_1(x_K) + g_0(x_K))$

viii-b) $$F = \begin{bmatrix} 0 & I \\ A_1 & A_2 \end{bmatrix} \quad \in BL(R^{2K})$$

$$A_1 = -C_3^{-1} C_1 = -\frac{c_1}{c_3} I = \frac{-\gamma_2}{\gamma_1 \delta_t + \gamma_2} I \qquad \in \; BL(R^K)$$

$$A_2 = C_3^{-1}(A - C_2) = \frac{1}{c_3}(A - c_2 I) =$$

$$= \begin{bmatrix} a_2' & a_3' & & & \\ a_1' & a_2' & a_3' & & \\ & \ddots & \ddots & \ddots & \\ & & a_1' & a_2' & a_3' \\ & & & a_1' & a_2' \end{bmatrix} \quad \in \; BL(R^K)$$

where:

$$a_1' = \frac{a_1}{c_3} = \alpha_2'$$

$$a_2' = \frac{a_2 - c_2}{c_3} = -\alpha_1' - 2\alpha_2' - \frac{c_2}{c_3} =$$

$$= 2 - (\alpha_1' + 2\alpha_2' + \delta_t \frac{\gamma_0 \; \delta_t + \gamma_1}{\gamma_1 \; \delta_t + \gamma_2})$$

$$a_3' = \frac{a_3}{c_3} = \alpha_1' + \alpha_2'$$

with the coefficients α_1' and α_2' defined as follows:

$$\alpha_1' = \frac{1}{c_3} \frac{\alpha_1}{\delta_x} = \frac{1}{\gamma_1 \; \delta_t + \gamma_2} \frac{\delta_t^2}{\delta_x} \alpha_1$$

$$\alpha_2' = \frac{1}{c_3} \frac{\alpha_2}{\delta_x^2} = \frac{1}{\gamma_1 \; \delta_t + \gamma_2} \frac{\delta_t^2}{\delta_x^2} \alpha_2$$

viii-b) $$F = \begin{bmatrix} 0 & I \\ A_1 & A_2 \end{bmatrix} \in BL(R^{2K})$$

$$A_1 = -C_3^{-1}\, C_1 = -\frac{c_1}{c_3}\, I = \frac{-\gamma_2}{\gamma_1 \delta_t + \gamma_2}\, I \qquad \in \; BL(R^K)$$

$$A_2 = C_3^{-1}(A - C_2) = \frac{1}{c_3}\,(A - c_2\, I) =$$

$$= \begin{bmatrix} a_2' & a_3' & & & \\ a_1' & a_2' & a_3' & & \\ & \ddots & \ddots & \ddots & \\ & & a_1' & a_2' & a_3' \\ & & & a_1' & a_2' \end{bmatrix} \in BL(R^K)$$

where:

$$a_1' = \frac{a_1}{c_3} = \alpha_2'$$

$$a_2' = \frac{a_2 - c_2}{c_3} = -\alpha_1' - 2\alpha_2' - \frac{c_2}{c_3} =$$

$$= 2 - (\alpha_1' + 2\,\alpha_2' + \delta_t\, \frac{\gamma_0\, \delta_t + \gamma_1}{\gamma_1\, \delta_t + \gamma_2})$$

$$a_3' = \frac{a_3}{c_3} = \alpha_1' + \alpha_2'$$

with the coefficients α_1' and α_2' defined as follows:

$$\alpha_1' = \frac{1}{c_3}\, \frac{\alpha_1}{\delta_x} = \frac{1}{\gamma_1\, \delta_t + \gamma_2}\, \frac{\delta_t^2}{\delta_x}\, \alpha_1$$

$$\alpha_2' = \frac{1}{c_3}\, \frac{\alpha_2}{\delta_x^2} = \frac{1}{\gamma_1\, \delta_t + \gamma_2}\, \frac{\delta_t^2}{\delta_x^2}\, \alpha_2$$

ix-b) $G = \begin{bmatrix} 0 \\ B_2 \end{bmatrix} \; : \; R^K \rightarrow R^{2K}$

$$B_2 = C_3^{-1} B = \frac{1}{c_3} B = \frac{\delta_t^2}{\gamma_1 \delta_t + \gamma_2} B \qquad \in \quad BL(R^K)$$

x-b) $H = h \begin{bmatrix} 0 & I \end{bmatrix} \; : \; R^{2K} \longrightarrow R^K$

Stochastic approximation algorithms for identification:

1st case: The original procedure:

(SA-1): $$\underline{a}(n+1) = \frac{1-\lambda(n)}{1-\mu(n)} \left[I - \mu(n) \, Q(n+1) \, \underline{z}_{k_o}(n+N_t) \, \underline{z}^*_{k_o}(n+N_t) \right] \underline{a}(n) + $$
$$+ \lambda(n) \, Q(n+1) \left[z_{k_o}(\tau_{n+N_t}) \, \underline{z}_{k_o}(n+N_t) + \underline{\varepsilon} \right] \quad : \quad \text{in } R^3$$

with $Q(n)$ in $BL(R^3)$ given by

(SA-2): $$Q(n+1) = \frac{1}{1-\mu(n)} \left[Q(n) - \frac{\mu(n) \, Q(n) \, \underline{z}_{k_o}(n+N_t) \, \underline{z}^*_{k_o}(n+N_t) \, Q(n)}{1-\mu(n) + \mu(n) \, \underline{z}^*_{k_o}(n+N_t) \, Q(n) \, \underline{z}_{k_o}(n+N_t)} \right]$$

where:

i) $n \in \bar{T} = \{0,1,2,\ldots\}$

ii) N_t : finite transient time as in (A - 4.4)

iii) k_o : any fixed point in $\overset{o}{S} = \{2,\ldots,K-1\}$

iv) $\lambda(n) \in (0,1) \quad ; \quad \sum_{n=0}^{\infty} \lambda(n) = \infty \quad ; \quad \sum_{n=0}^{\infty} \lambda^2(n) < \infty$

v) $\mu(n) \in (0,1) \quad ; \quad \sum_{n=0}^{\infty} \mu(n) = \infty \quad ; \quad \sum_{n=0}^{\infty} \mu^2(n) < \infty$

vi) $$z_{k_o}(\tau_n) = \begin{cases} c_1\, z_{k_o}(n) + c_2\, z_{k_o}(n+1) & ; \quad \underline{N=1} \\ c_1\, z_{k_o}(n-1) + c_2\, z_{k_o}(n) + c_3\, z_{k_o}(n+1) & ; \quad \underline{N=2} \end{cases}$$

vii) $\underline{z}_{k_o}(n) = J_{k_o}\, \underline{z}(n) = (z_{k_o-1}, z_{k_o}(n), z_{k_o+1}(n))$: in R^3

viii) $$\varepsilon = -d^2 \eta_v^2\, \gamma_0 - d^2 \begin{cases} c_1\, \delta_v(0) + c_2\, \delta_v(1) & ; \quad \underline{N=1} \\ c_1\, \delta_v(1) + c_2\, \delta_v(0) + c_3\, \delta_v(1) & ; \quad \underline{N=2} \end{cases}$$

ix) $\underline{\varepsilon} = \varepsilon(1,1,1) \in R^3$

x) $\underline{a}(0)$: a second-order random vector in R^3, independent of $\{\underline{z}_{k_o}(n+N_t)\ ;\ n \in \bar{T}\}$.

xi) $Q(0)$: a second-order symmetric positive definite random matrix in $BL(R^3)$, independent of $\{\underline{z}_{k_o}(n+N_t)\, \underline{z}^*_{k_o}(n+N_t)\ ;\ n \in \bar{T}\}$.

<u>2nd case</u>: An equivalent procedure:

(SA-1): $$\underline{a}'(n+1) = \frac{1-\lambda(n)}{1-\mu(n)} \left[I - \mu(n)\, Q(n+1)\, \underline{z}_{k_o}(n+N_t)\, \underline{z}^*_{k_o}(n+N_t) \right] \underline{a}'(n) + \lambda(n)\, Q(n+1) \left[z_{k_o}(\tau'_{n+N_t})\, \underline{z}_{k_o}(n+N_t) + \underline{\varepsilon}' \right] : \quad \text{in } R^3$$

with

vi') $$z_{k_o}(\tau'_n) = \begin{cases} z_{k_o}(n+1) & ; \quad \underline{N=1} \\ \dfrac{c_1}{c_3}\, z_{k_o}(n-1) + z_{k_o}(n+1) & ; \quad \underline{N=2} \end{cases}$$

ix') $\underline{\varepsilon}' = \begin{cases} \frac{1}{c_2}\,\underline{\varepsilon}\ ; & \underline{N=1} \\ \frac{1}{c_3}\,\underline{\varepsilon}\ ; & \underline{N=2} \end{cases}$

x') $\underline{a}'(0)$: a second-order random vector in R^3, independent of $\{\underline{z}_k(n+N_t)\ ;\ n \in \bar{T}\}$.

In both cases the algorithms converge with probability one:

$$\underline{a}(n) \xrightarrow[n\to\infty]{w.p.1} \underline{a} = (a_1\,,a_2\,,a_3) \in R^3$$

$$\underline{a}'(n) \xrightarrow[n\to\infty]{w.p.1} \underline{a}' = (a_1'\,,a_2'\,,a_3') \in R^3$$

Recovering the original parameters:

1st case: Using the original procedure:

$$\underline{a} = R\,\underline{\alpha}$$

where:

i) $\underline{a} = (a_1\,,a_2\,,a_3) \in R^3$

ii) $\underline{\alpha} = (\alpha_1\,,\alpha_2) \in R^2$

iii) $R = \left[r_{im}\right] = \begin{bmatrix} r_{-11} & r_{-12} \\ r_{01} & r_{02} \\ r_{11} & r_{12} \end{bmatrix} = \frac{1}{\delta_x^2}\begin{bmatrix} 0 & 1 \\ -\delta_x & -2 \\ \delta_x & 1 \end{bmatrix} : R^2 \to R^3$

The least-squares estimate is given by

$$\hat{\underline{\alpha}}(n) = (R^* R)^{-1} R^*\,\underline{a}_k(n) \xrightarrow[n\to\infty]{w.p.1} \underline{\alpha}$$

where:

iv) $\underline{\hat{\alpha}}(n) = (\hat{\alpha}_1(n), \hat{\alpha}_2(n))$: in R^2

v) $\underline{a}(n) = (a_1(n), a_2(n), a_3(n))$: in R^3

vi) $$(R^* R)^{-1} R^* = \frac{\delta_x}{3} \begin{bmatrix} -3 & 0 & 1 \\ 2\delta_x & -\delta_x & -\delta_x \end{bmatrix} \quad : \quad R^3 \longrightarrow R^2$$

So we get:

$$\hat{\alpha}_1(n) = \delta_x \left[- a_1(n) + a_3(n) \right]$$

$$\hat{\alpha}_2(n) = \frac{\delta_x^2}{3} \left[2\, a_1(n) - a_2(n) - a_3(n) \right]$$

<u>2nd case</u>: Using the equivalent procedure:

$$\underline{a}' = R' \, \underline{\alpha}' - \frac{c^*_{N+1}}{c_{N+1}} \, \underline{e}_2$$

where:

i) $\underline{a}' = (a'_1, a'_2, a'_3) \in R^3$

ii) $\underline{\alpha}' = (\alpha'_1, \alpha'_2) \in R^2$

<u>N = 1</u>: $$\begin{cases} \alpha'_1 = \dfrac{1}{c_2} \dfrac{\alpha_1}{\delta_x} = \dfrac{1}{\delta_1} \dfrac{\delta_t}{\delta_x} \alpha_1 \\[2ex] \alpha'_2 = \dfrac{1}{c_2} \dfrac{\alpha_2}{\delta_x^2} = \dfrac{1}{\delta_1} \dfrac{\delta_t}{\delta_x^2} \alpha_2 \end{cases}$$

$$\underline{N=2}: \begin{cases} \alpha'_1 = \dfrac{1}{c_3}\,\dfrac{\alpha_1}{\delta_x} = \dfrac{1}{\gamma_1\ \delta_t + \gamma_2}\,\dfrac{\delta_t^2}{\delta_x}\,\alpha_1 \\[2ex] \alpha'_2 = \dfrac{1}{c_3}\,\dfrac{\alpha_2}{\delta_x^2} = \dfrac{1}{\gamma_1\ \delta_t + \gamma_2}\,\dfrac{\delta_t^2}{\delta_x^2}\,\alpha_2 \end{cases}$$

iii) $R' = \left[r'_{im}\right] = \left[r_{im}\ \delta_x^m\right] = \begin{bmatrix} 0 & 1 \\ -1 & -2 \\ 1 & 1 \end{bmatrix} \ : \ R^2 \longrightarrow R^3$

iv) $\underline{e}_2 = (0\,,1\,,0) \in R^3$

v) $$\frac{c^{*}_{N+1}}{c_{N+1}} = \begin{cases} \dfrac{c_1}{c_2} = \delta_t\,\dfrac{\gamma_0}{\gamma_1} - 1 \quad ; \quad \underline{N=1} \\[2ex] \dfrac{c_2}{c_3} = \delta_t\,\dfrac{\gamma_0}{\gamma_1}\,\dfrac{\delta_t + \gamma_1}{\delta_t + \gamma_2} - 2 \quad ; \quad \underline{N=2} \end{cases}$$

The least-squares estimate is given by

$$\underline{\hat{\alpha}}'(n) = (R'^{*}R')^{-1}\,R'^{*}\left[\underline{a}'(n) + \frac{c^{*}_{N+1}}{c_{N+1}}\,\underline{e}_2\right] \xrightarrow[n\to\infty]{\text{w.p.1}} \underline{\alpha}'$$

where:

vi) $\underline{\hat{\alpha}}'(n) = (\hat{\alpha}'_1(n)\,,\hat{\alpha}'_2(n))$: in R^2

vii) $\underline{a}'(n) = (a'_1(n)\,,\,a'_2(n)\,,\,a'_3(n))$: in R^3

viii) $(R'^{*}R')^{-1}\,R'^{*} = \dfrac{1}{3}\begin{bmatrix} 3 & 0 & 3 \\ 2 & -1 & -1 \end{bmatrix} \ : \ R^3 \longrightarrow R^2$

So we get:

$$\underline{N=1}: \begin{cases} \hat{\alpha}'_1(n) = -a'_1(n) + a'_3(n) \\[1ex] \hat{\alpha}'_2(n) = \dfrac{1}{3}\left[2\,a'_1(n) - a'_2(n) - a'_3(n) + 1 - \delta_t\,\dfrac{\gamma_0}{\gamma_1}\right] \end{cases}$$

$$
\underline{N=2}: \begin{cases} \hat{\alpha}'_1(n) = -a'_1(n) + a'_3(n) \\ \hat{\alpha}'_2(n) = \dfrac{1}{3} \left[2\, a'_1(n) - a'_2(n) - a'_3(n) + 2 - \delta_t \, \dfrac{\gamma_0 \ \delta_t + \gamma_1}{\gamma_1 \ \delta_t + \gamma_2} \right] \end{cases}
$$

Remark: If $\alpha_1 = 0$ is known "a priori", we get the least-squares estimate for α_2 as follows:

1st case:

$$
\underline{a} = R\,\alpha_2
$$

$$
R = \frac{1}{\delta_x^2} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \quad : \quad R^1 \longrightarrow R^3
$$

So:

$$
\hat{\alpha}_2(n) = \frac{\delta_x^2}{6} \left[a_1(n) - 2\, a_2(n) + a_3(n) \right]
$$

2nd case:

$$
\underline{a}' = R'\,\alpha'_2 - \frac{c^{*}_{N+1}}{c_{N+1}}\,\underline{e}_2
$$

$$
R' = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \quad : \quad R^1 \longrightarrow R^3
$$

So:

$\underline{N=1}$: $$\hat{\alpha}'_2(n) = \frac{1}{6} \left[a'_1(n) - 2\, a'_2(n) + a'_3(n) + 2 - 2\,\delta_t \, \frac{\gamma_0}{\gamma_1} \right]$$

$\underline{N=2}$: $$\hat{\alpha}'_2(n) = \frac{1}{6} \left[a'_1(n) - 2\, a'_2(n) + a'_3(n) + 4 - 2\,\delta_t \, \frac{\gamma_0 \ \delta_t + \gamma_1}{\gamma_1 \ \delta_t + \gamma_2} \right]$$

A block diagram: Figure 8 shows a block diagram for system simulation and identification. The DPS simulation is carried out by using the equivalent discrete-time LPS. The parameter vector $\underline{a}$ is identified on-line, through noisy observations $\underline{z}(n+N_t)$, via stochastic approximation algorithms given in (SA-1) and (SA-2). The original parameters $\underline{\alpha} = (\alpha_1, \alpha_2)$ are recovered from $\underline{a}(n)$ by means of a simplo least-squares.

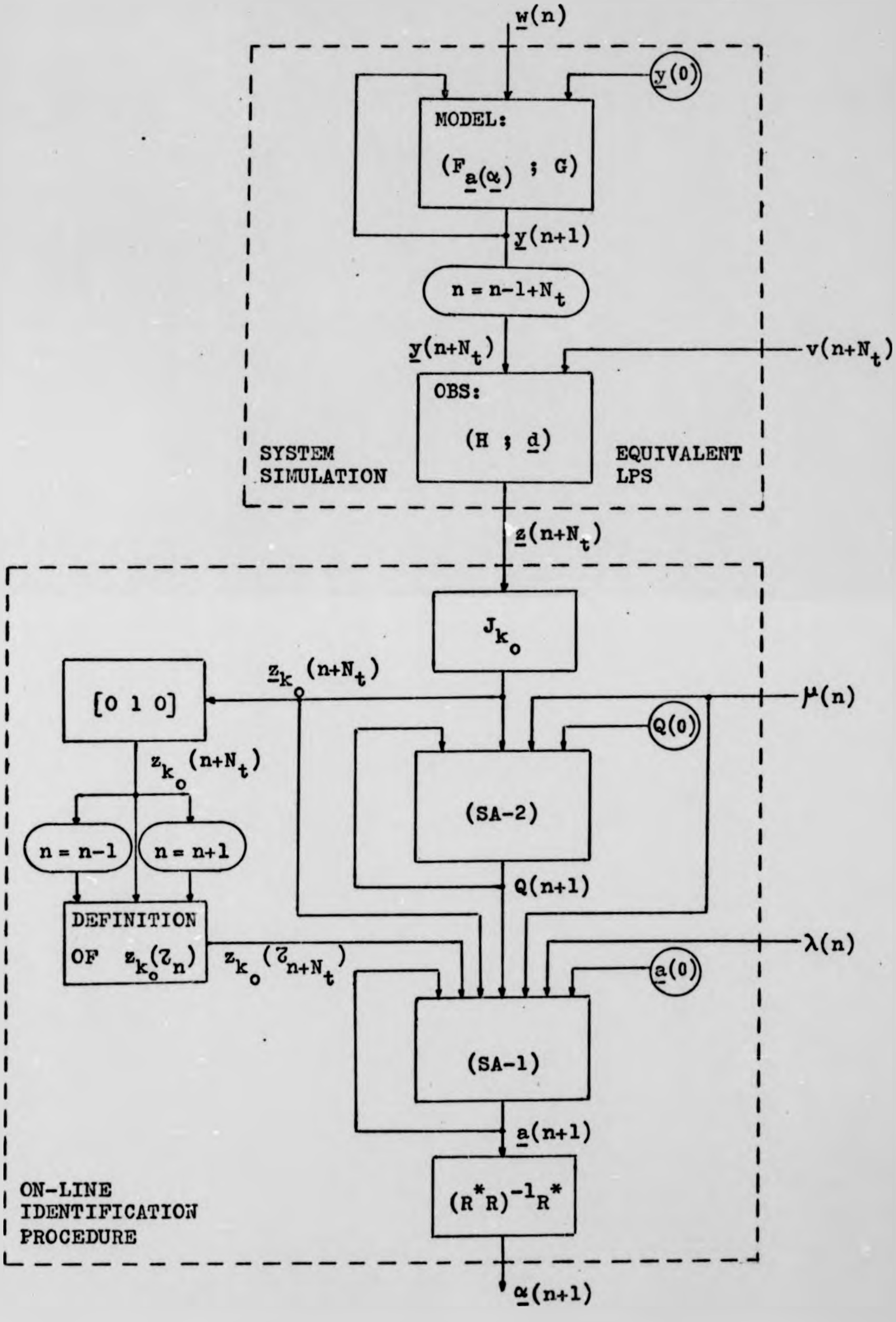

Fig. 8: SYSTEM SIMULATION AND IDENTIFICATION.

5.2 - GENERAL ASSUMPTIONS FOR EXAMPLES

Examples illustrating this DPS identification method will be presented in sections 5.3 through 5.5. The following assumptions were made when performing those examples.

Simulation: The original DPS was simulated by using the equivalent discrete-time LPS, where the spatial domain

$$X = (0,\pi)$$

(i.e., $\ell = \pi$) was discretized with

$$\delta_x = \pi/6$$

So,

$$K = 5 \rightarrow \overset{\circ}{S} = \{2,3,4\} \subset S = \{1,2,3,4,5\}$$

The M+1=3 observation points in X (i.e., x_{k_o-1}, x_{k_o} and x_{k_o+1} where k_o is any fixed point in $\overset{\circ}{S}$), the constants h and d, the time sampling rate δ_t, the space-varying input parameter $\beta(x)$, and the initial functions $g_0(x)$ and $g_1(x)$ were chosen as follows:

$$\text{i)}\quad k_o = 3 \rightarrow \begin{cases} x_{k_o-1} = (k_o - 1)\,\delta_x = \pi/3 \\ x_{k_o} = k_o\,\delta_x = \pi/2 \\ x_{k_o+1} = (k_o + 1)\,\delta_x = 2\pi/3 \end{cases}$$

$$\text{ii)}\quad h = d = 1$$

$$\text{iii)}\quad \delta_t = \begin{cases} 1/2 \; ; & \text{if } N = 1 \\ \sqrt{1/2} \; ; & \text{if } N = 2 \end{cases}$$

5.2 - GENERAL ASSUMPTIONS FOR EXAMPLES

Examples illustrating this DPS identification method will be presented in sections 5.3 through 5.5. The following assumptions were made when performing those examples.

Simulation: The original DPS was simulated by using the equivalent discrete-time LPS, where the spatial domain

$$X = (0,\pi)$$

(i.e., $\ell = \pi$) was discretized with

$$\delta_x = \pi/6$$

So,

$$K = 5 \rightarrow \mathring{S} = \{2,3,4\} \subset S = \{1,2,3,4,5\}$$

The M+1=3 observation points in X (i.e., x_{k_o-1}, x_{k_o} and x_{k_o+1} where k_o is any fixed point in $\mathring{S}$), the constants h and d, the time sampling rate δ_t, the space-varying input parameter $\beta(x)$, and the initial functions $g_0(x)$ and $g_1(x)$ were chosen as follows:

i) $$k_o = 3 \rightarrow \begin{cases} x_{k_o-1} = (k_o - 1)\,\delta_x = \pi/3 \\ x_{k_o} = k_o\,\delta_x = \pi/2 \\ x_{k_o+1} = (k_o + 1)\,\delta_x = 2\pi/3 \end{cases}$$

ii) $h = d = 1$

iii) $$\delta_t = \begin{cases} 1/2 \; ; & \text{if } N = 1 \\ \sqrt{1/2} \; ; & \text{if } N = 2 \end{cases}$$

iv) $\beta(x) = 2\sin(x)$; $x \in (0,\pi)$

v) $g_0(x) = g_1(x) = 0$; $x \in [0,\pi]$

Input disturbance: $\{w_k(n) \; ; \; n \in T'\}$ was chosen to be uniformly distributed in $(-\sqrt{3}\,\sigma_w , \sqrt{3}\,\sigma_w)$ and uncorrelated for all $k \in S$, such that:

$$\sigma_w^2 = E\{w_k^2(n)\} \; ; \qquad \text{for all } k \in S$$

$$C_w = \sigma_w^2 \, I \quad \in \quad BL(R^5)$$

with $\sigma_w = 1/2$ (the example in section 5.3 also consider the case with $\sigma_w = 1$).

Observation noise: $\{v(n) \; ; \; n \in T\}$ was chosen to be uniformly distributed in $(-\sqrt{3}\,\sigma_v , \sqrt{3}\,\sigma_v)$ such that:

$$E\{v(n)\} = \eta_v = 0$$

$$E\{v(i)\,v(j)\} - \eta_v^2 = \sigma_v(i-j) = \sigma_v^2\,\delta(i-j)$$

with $\sigma_v = 1/4$.

Stochastic approximation algorithms (SA-1) and (SA-2): The identification was carried out by using the equivalent procedure; that is, the algorithm in (SA-1) was

$$\underline{a}'(n) = (a_1'(n) , a_2'(n) , a_3'(n)) \xrightarrow[n \to \infty]{w.p.1} \underline{a}' = (a_1' , a_2' , a_3') \in R^3$$

The following situation was assumed:

i) $N_t = 100$

ii) $\lambda(n) = \dfrac{1}{n+2}$

iii) $\mu(n) = \dfrac{1}{n + 3/2}$

iv) $\underline{a}'(0) = \begin{cases} (1/2\ ,\ 0\ ,\ 1/2) \in R^3\ ; & \text{when } N = 1 \\ (1\ ,\ 0\ ,\ 1) \in R^3\ ; & \text{when } N = 2 \end{cases}$

v) $Q(0) = I \in BL(R^3)$

5.3 - PARABOLIC PDE WITH ONE PARAMETER:

Our first example considers the identification of a single parameter α_2 appearing in the "heat equation":

$$\frac{\partial}{\partial t} u(x,t) = \alpha_2 \frac{\partial^2}{\partial x^2} u(x,t) + 2 \sin(x)\, w_x(t)$$

(i.e: $N = 1$, $\gamma_0 = 0$, $\gamma_1 = 1$ and $\alpha_1 = 0$ known "a priori"). The simulation was carried out by using

$$\alpha_2 = \frac{1}{2}\left(\frac{\pi}{6}\right)^2 \longrightarrow \alpha_2' = \frac{1}{4}$$

and so, the constant vector to be identified by (SA-1) is

$$\underline{a}' = (a_1', a_2', a_3') = (\alpha_2', 1 - 2\alpha_2', \alpha_2') .$$

Figures 9 and 10 show the performance of the identification procedure, where

$$\hat{\alpha}_2'(n) = \frac{1}{6}\left[a_1'(n) + 2\,(1 - a_2'(n)) + a_3'(n)\right]$$

$$\tilde{a}'(n) = \|\underline{a}'(n) - \underline{a}'\|^2$$

Two cases were considered:

i) $\sigma_w = 1/2$

ii) $\sigma_w = 1$

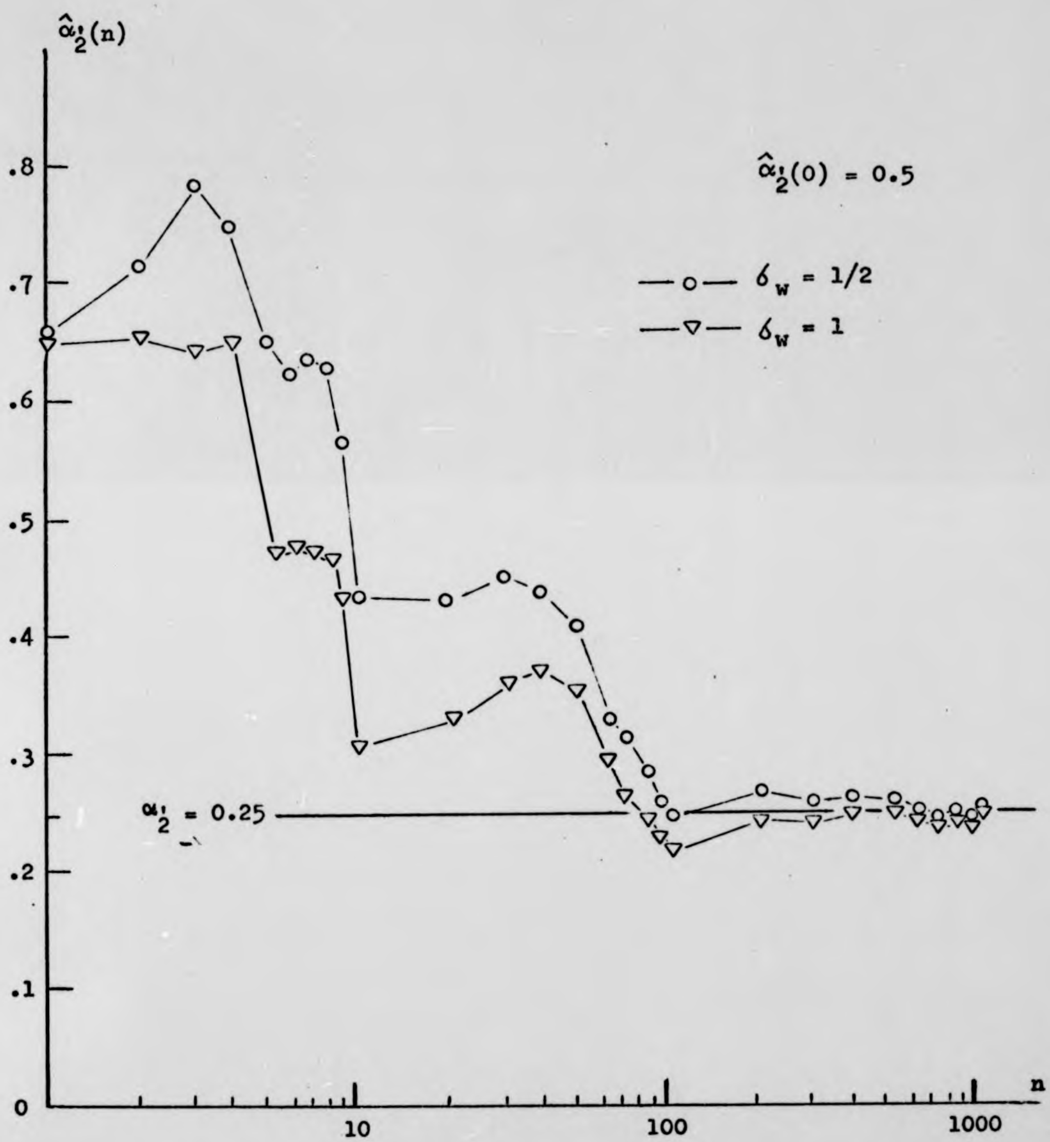

Fig. 9: Estimate Performance for α'_2.

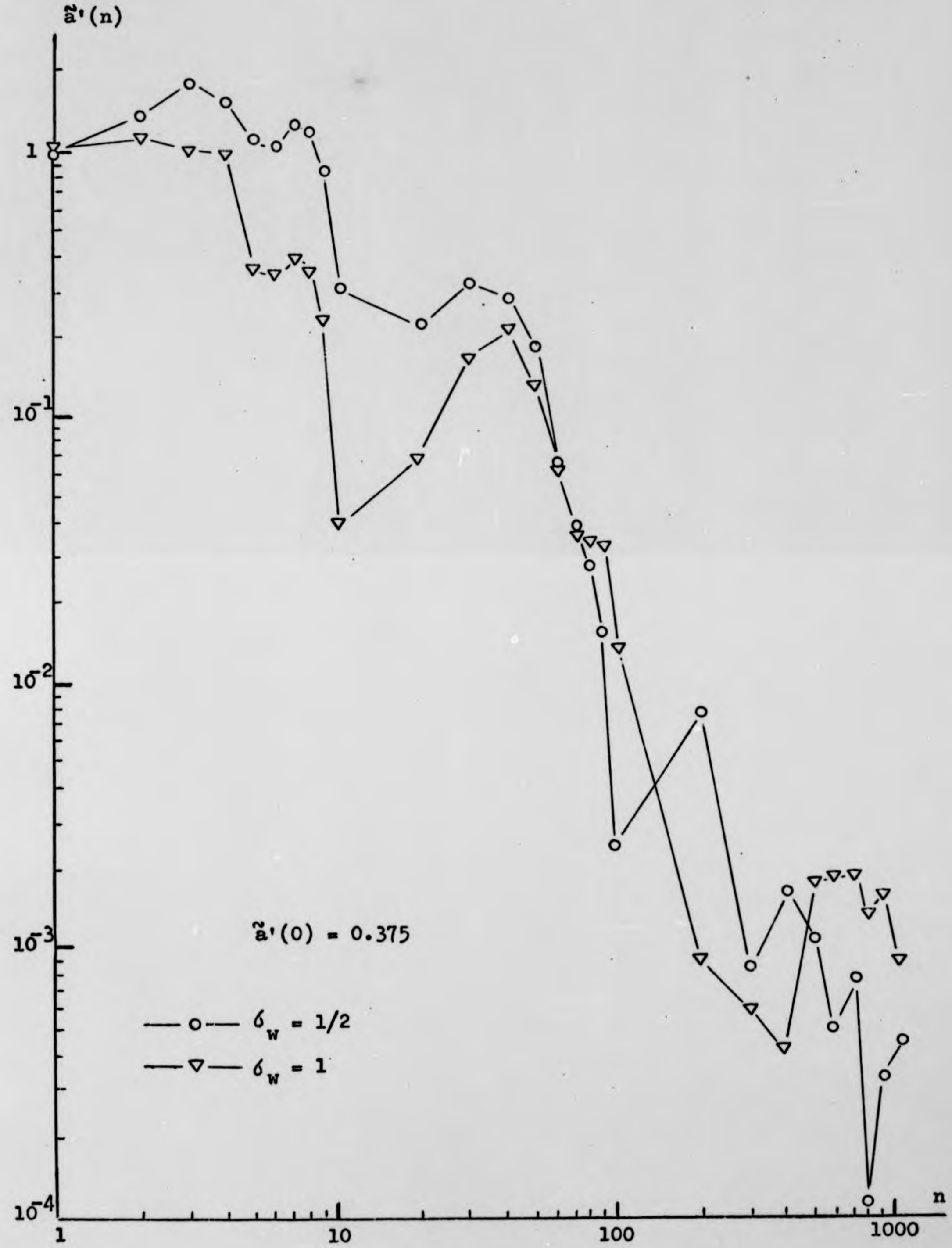

Fig. 10: Error Performance.

5.4 - PARABOLIC PDE WITH TWO PARAMETERS:

Now we consider the identification of α_1 and α_2 appearing in a parabolic equation, as follows:

$$\frac{\partial}{\partial t} u(x,t) = \alpha_1 \frac{\partial}{\partial x} u(x,t) + \alpha_2 \frac{\partial^2}{\partial x^2} u(x,t) + 2\sin(x)\, w_x(t)$$

(i.e.: $N = 1$, $\gamma_0 = 0$, $\gamma_1 = 1$). For simulating the DPS we assumed

$$\alpha_2 = \frac{1}{2}\left(\frac{\pi}{6}\right)^2 \longrightarrow \alpha_2' = \frac{1}{4}$$

and two cases were considered (the first one representing a model with extraneous terms):

i) $\alpha_1 = 0 \longrightarrow \alpha_1' = 0$

ii) $\alpha_1 = -\frac{\pi}{30} \longrightarrow \alpha_1' = -\frac{1}{10}$

In this way, the constant vector to be identified by (SA-1) is:

$$\underline{a}' = (a_1', a_2', a_3') = (\alpha_2', 1-(\alpha_1'+2\alpha_2'), \alpha_1'+\alpha_2').$$

The performance of the identification procedure is shown on figures 11 - 13, where

$$\hat{\alpha}_1'(n) = -a_1'(n) + a_3'(n)$$

$$\hat{\alpha}_2'(n) = \frac{1}{3}\left[2\, a_1'(n) - a_2'(n) - a_3'(n) + 1\right]$$

$$\tilde{a}'(n) = \|\underline{a}'(n) - \underline{a}'\|^2$$

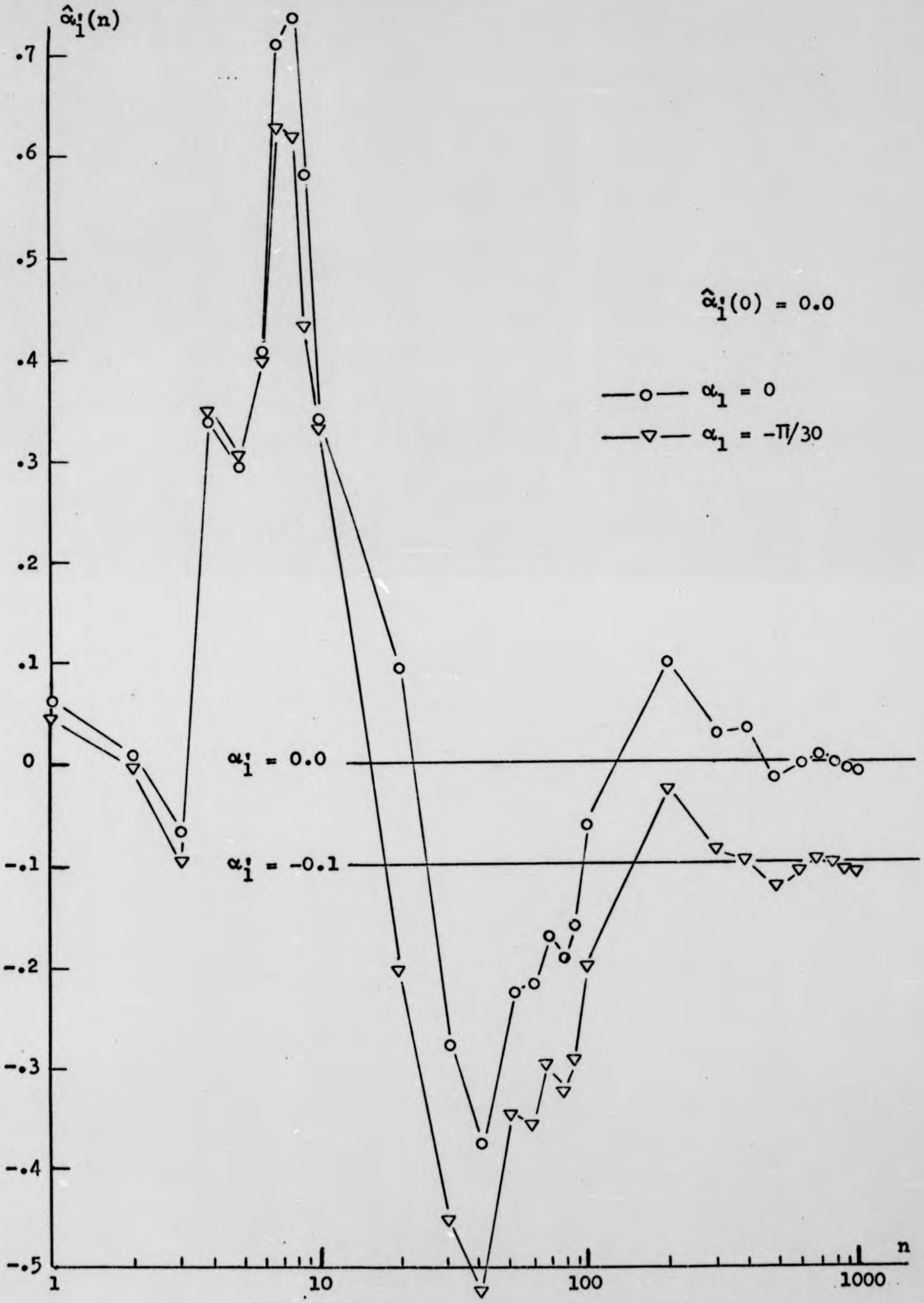

Fig. 11: Estimate Performance for α_1'.

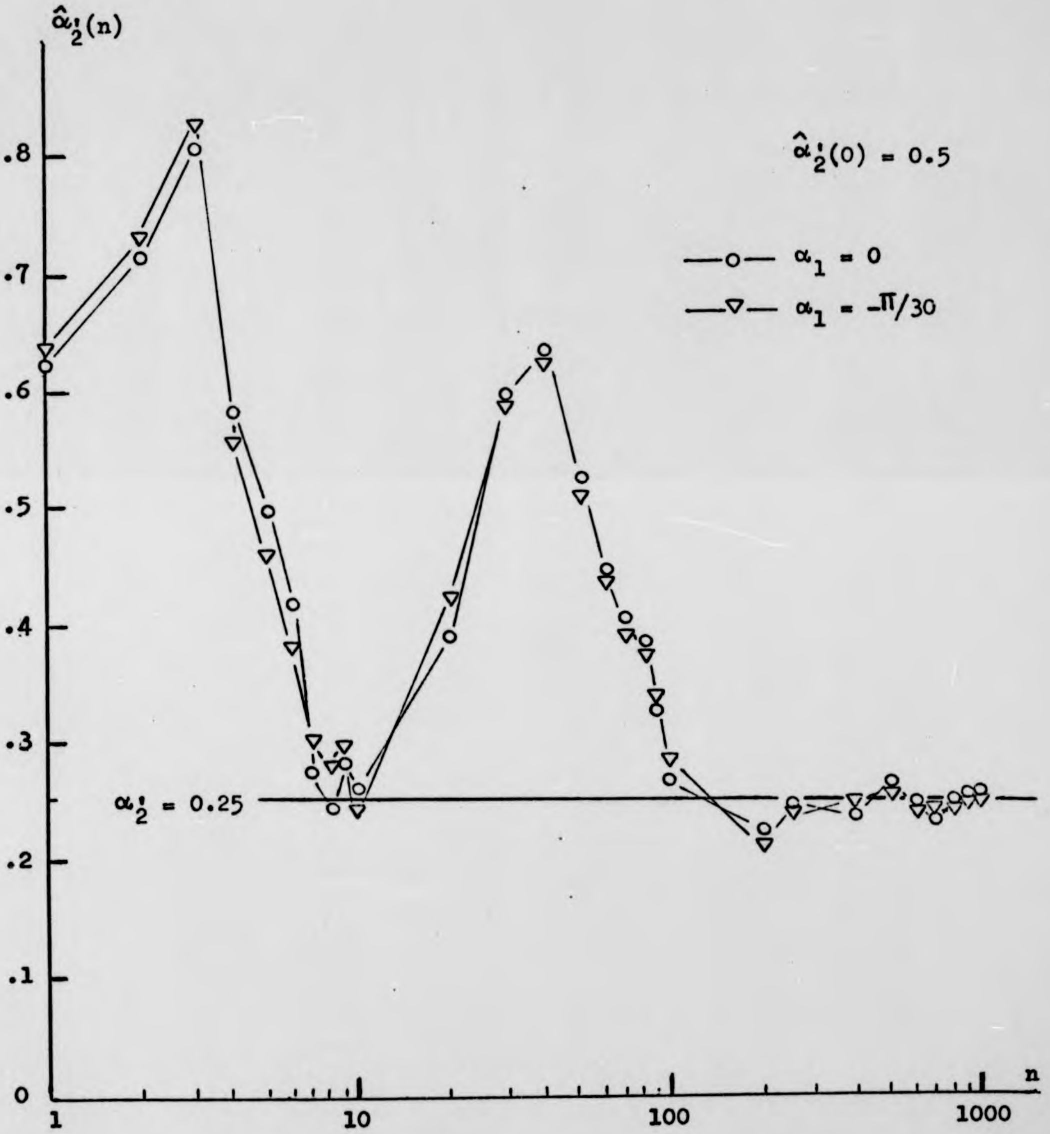

Fig. 12: Estimate Performance for α_2'.

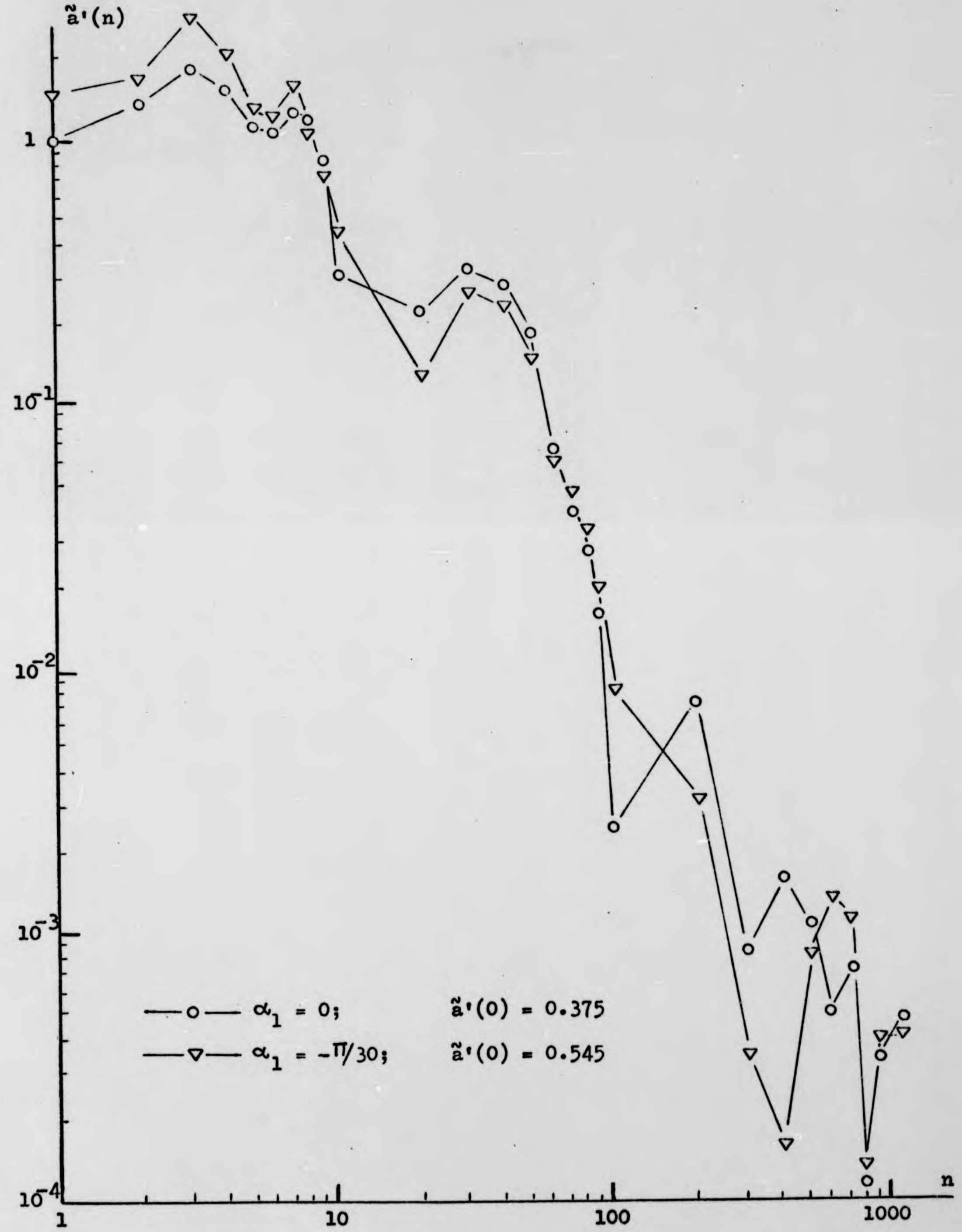

Fig. 13: Error Performance.

5.5 - HYPERBOLIC PDE WITH TWO PARAMETERS:

Finally we present an example illustrating the identification of α_1 and α_2 appearing in a hyperbolic equation:

$$\frac{\partial^2}{\partial t^2} u(x,t) = \alpha_1 \frac{\partial}{\partial x} u(x,t) + \alpha_2 \frac{\partial^2}{\partial x^2} u(x,t) + 2\sin(x)\; w_x(t)$$

(i.e.: $N = 2$, $\gamma_0 = \gamma_1 = 0$, $\gamma_2 = 1$). The simulation was carried out by assuming

$$\alpha_2 = \left(\frac{\pi}{6}\right)^2 \longrightarrow \alpha_2' = \frac{1}{2}$$

and two cases considered:

$$\text{i)}\quad \alpha_1 = \frac{\pi}{30} \longrightarrow \alpha_1' = \frac{1}{10}$$

$$\text{ii)}\quad \alpha_2 = -\frac{\pi}{30} \longrightarrow \alpha_1' = -\frac{1}{10}$$

So, the constant vector to be identified by (SA-1) is

$$\underline{a}' = (a_1', a_2', a_3') = (\alpha_2', 2-(\alpha_1'+2\alpha_2'), \alpha_1'+\alpha_2') .$$

Figures 14 - 16 shown the performance of the identification procedure, where

$$\hat{\alpha}_1'(n) = -a_1'(n) + a_3'(n)$$

$$\hat{\alpha}_2'(n) = \frac{1}{3}\left[2\,a_1'(n) - a_2'(n) - a_3'(n) + 2\right]$$

$$\tilde{a}'(n) = \|\underline{a}'(n) - \underline{a}\|^2$$

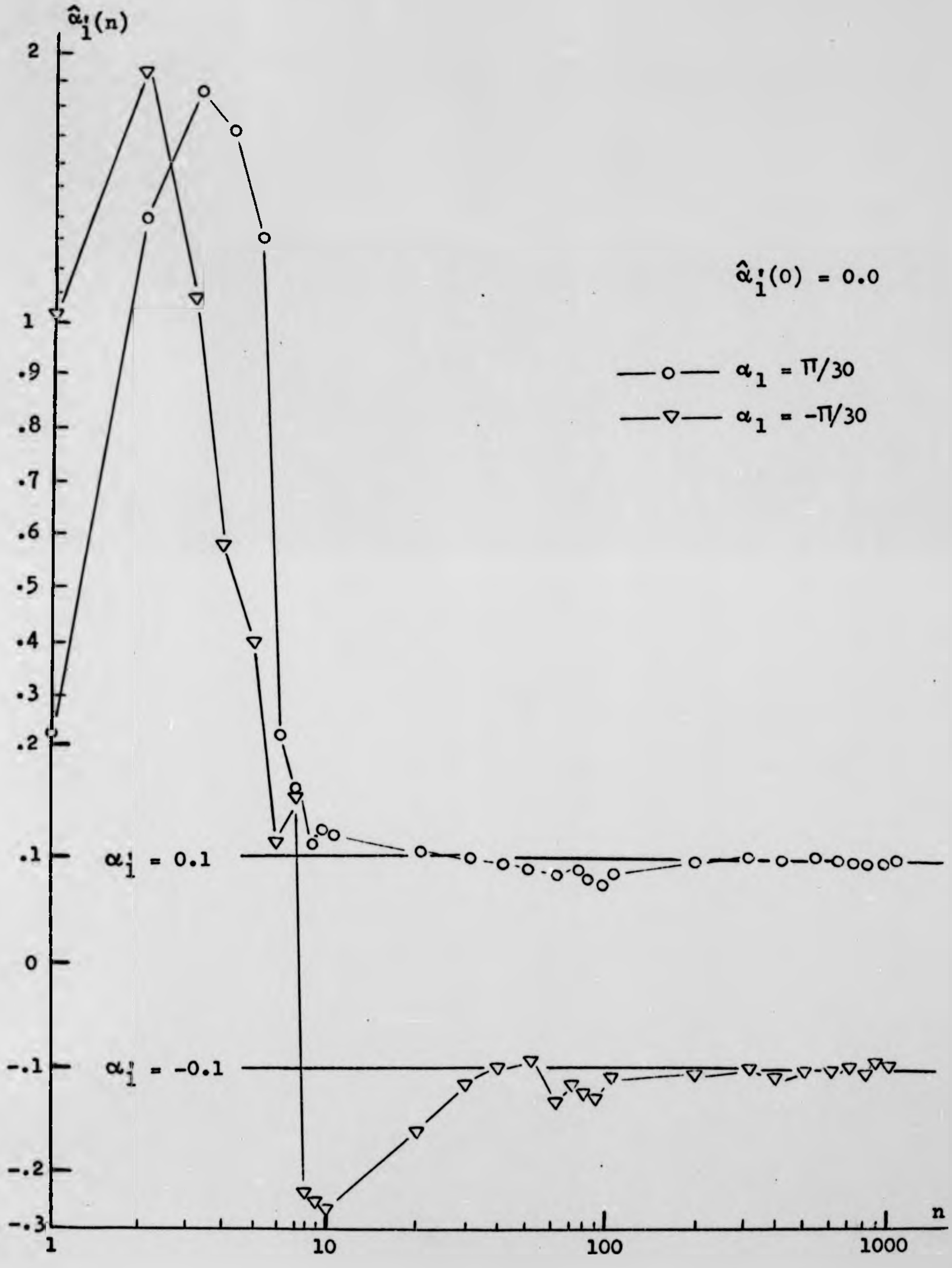

Fig. 14: Estimate Performance for α_1'.

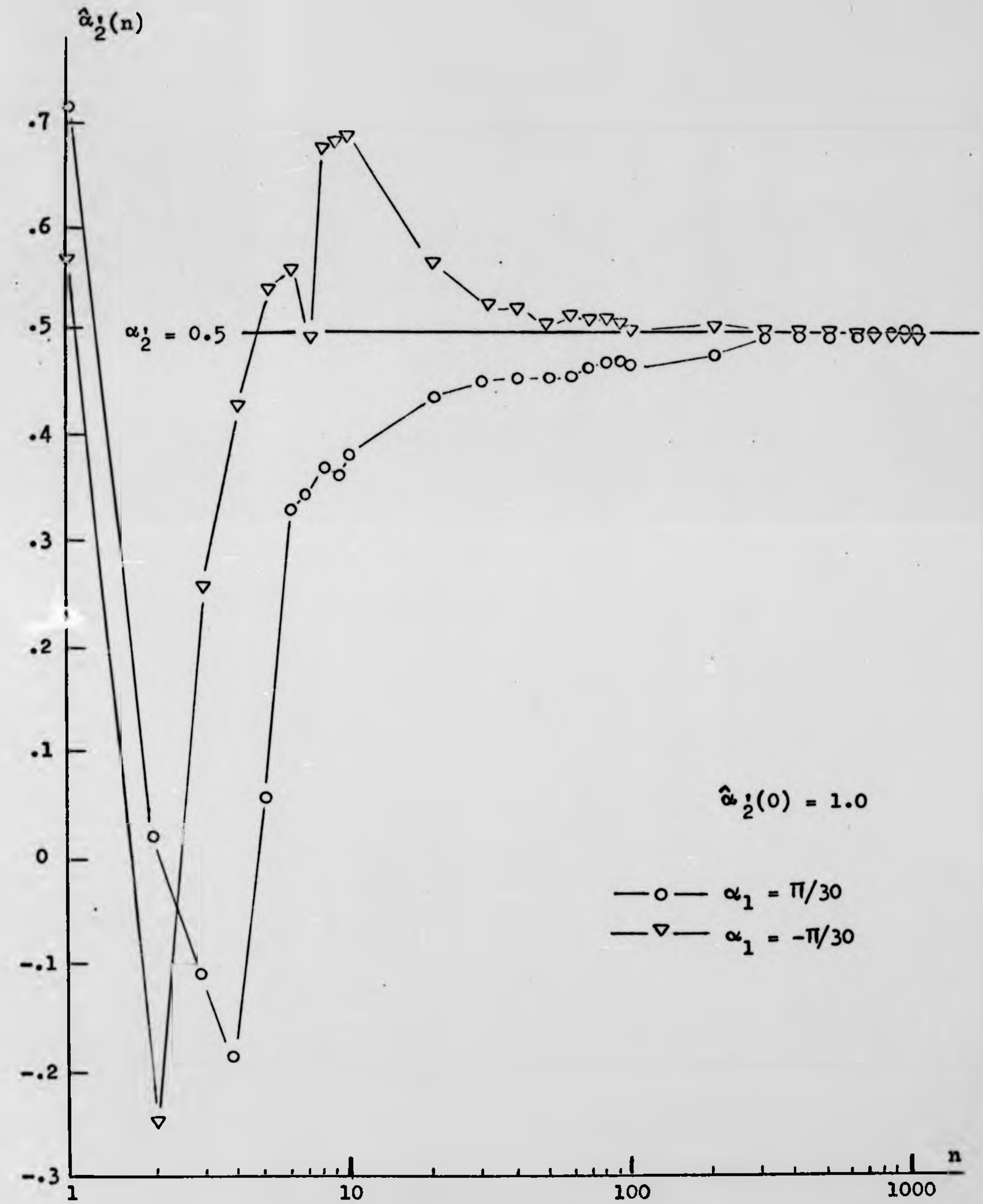

Fig. 15: Estimate Performance for α_2'.

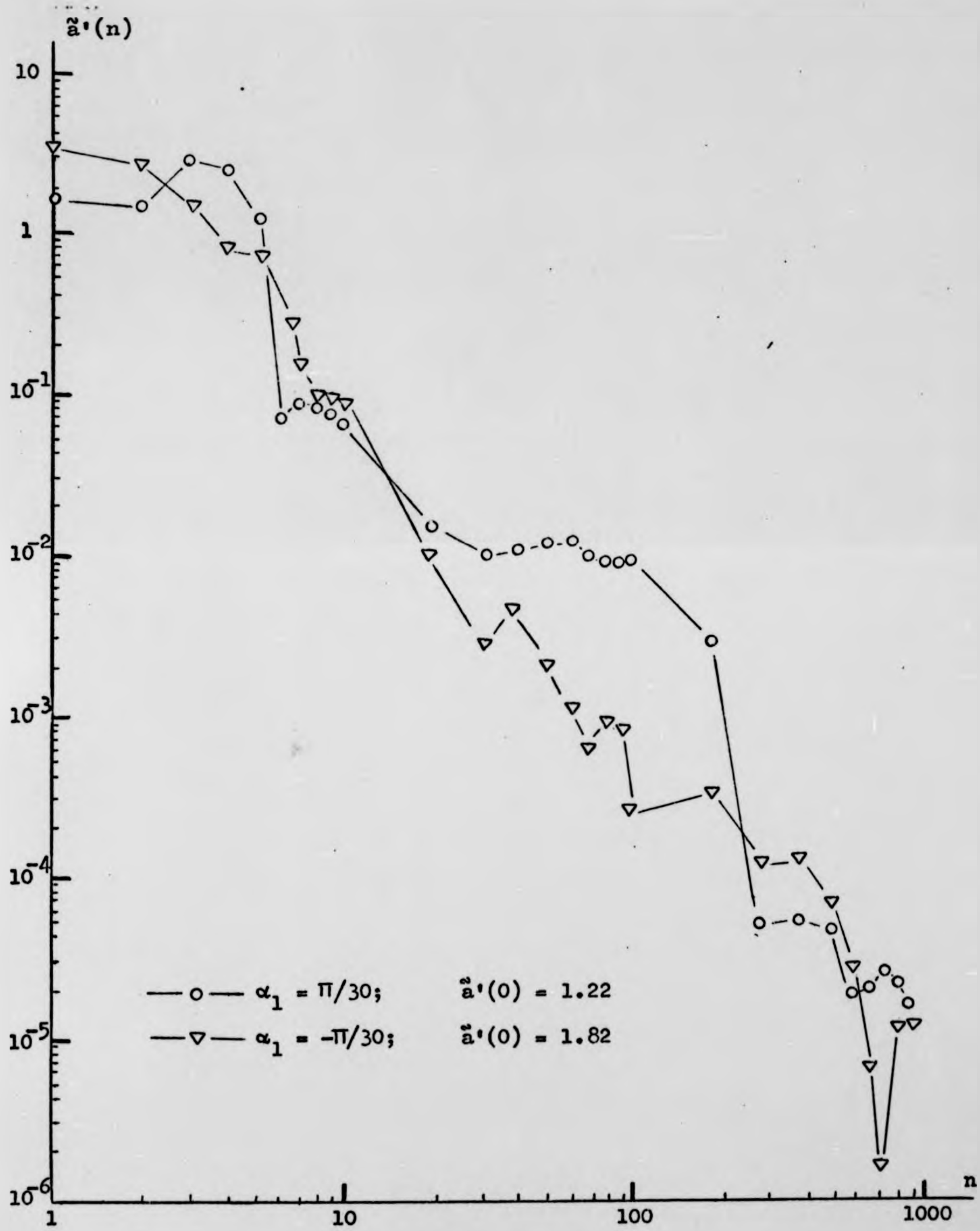

Fig. 16: Error Performance.

5.6 - CONCLUSIONS

At the end of chapter 2 we have surveyed and commented on the general DPS identification problem. In this final section we discuss the particular identification method introduced in this work.

The stochastic approximation theory was used to identify parameters in distributed systems operating in a stochastic environment. Some basic points which characterize the method are listed below.

1) CLASS OF MODELS: The models we have considered here are described by PDE and have the following properties:

   1-a) Linear, of order N in t and M in the scalar spatial variable x, with space-varying parameters.

   1-b) Extensions to multi-dimensional spatial domain ($x \in X \subset R^P$) were also considered in section 4.7.

   1-c) Although we have not considered cross-terms partial derivatives in our model, they can be treated using the same technique. But, in this case, the choice of which kind of approximation (backward, centered, or forward operators) must be decided for each model containing a particular type of cross-terms partial derivatives.

   1-d) The space-varying parameters to be identified are those multiplying spatial derivatives (parameters appearing in L_x^M, the spatial-differential operator), that is: $\{\alpha_m(x) \ ; \ m = 1,\dots,M\}$. Parameters appearing in L_t^N (i.e., $\{\gamma_m(x) \ ; \ m = 0,\dots,N\}$) are assumed to be known "a priori".

   1-e) In case of constant parameters to be identified, the identification procedure can be simplified as commented on page 114.

2) METHOD CLASSIFICATION: The method is classified as class Γ_2 (see section 2.2), and so it presents two stages: model approximation (stage I) and parametric estimation (stage II).

3) REDUCTION TO A FINITE-DIMENSIONAL STATE SPACE: In the first stage the method comprises two basic steps:

   3-a) An equivalent discrete-time LPS obtained by using higher order finite-differences (section 4.2).

   3-b) A fundamental observation equation given in (P - 4.1), which was possible thanks to the results obtained by finite-difference techniques (see remark 3 on page 115).

4) EXPLICIT PARAMETER: The parameters appearing in the discrete version (chapter 4, equation (5)) were placed in an explicit form (L - 4.1). In this way it was possible to use stochastic approximation algorithms as a straightforward identification procedure, rather than a simple searching scheme for finding estimates previously obtained by means of any other optimization technique.

5) INPUT DISTURBANCE: The input $\{\underline{w}(n)\}$ was taken to be a zero mean "white" random sequence with positive definite covariance matrix C_w, as in (A - 4.2). It is also possible to consider positive semi-definite covariance matrix (or even zero input), if the initial state is assumed to be random, as commented on page 103. It is worth to remember here that, in order to use our identification procedure (based on a straightforward applicability of stochastic approximation via explicit parameter lemma (L - 4.1)), we cannot have deterministic inputs and deterministic initial state together. But in this case the reduction to an equivalent discrete-time LPS (stage I), as developed in section 4.2, can still be used for identification

purposes by applying in stage II some other known technique (e.g., see surveys in LPS identification mentioned in chapter 1) for parametric estimation in LPS driven by deterministic inputs.

6) OBSERVATIONS: We have assumed that noisy observations are available at a finite number of discrete points equidistantly located in the spatial domain. In case of constant parameters just M+1 of those points are required, as remarked on page 114. Also, in some special cases, the measurements can be taken at pre-selected (not necessarily equidistant) observation points, as commented in [1].

7) OBSERVATION NOISE: The noise $\{v(n)\}$ corrupting the observations is assumed to be uncorrelated with the input disturbance as in (A - 4.2), and its statistics $\{\eta_v, \delta_v(0), \ldots, \delta_v(N)\}$ are supposed to be known.

8) BOUNDARY CONDITIONS: Altough the method has been developed using homogeneous BC, nonhomogeneous BC may also be considered (see remark on page 93), as well as random boundary conditions.

9) INITIAL CONDITIONS: Both, deterministic and random initial state, can be considered as commented before (see pages 83 and 103).

10) STOCHASTIC APPROXIMATION ALGORITHMS FOR IDENTIFICATION: The algorithms in (SA-1) and (SA-2) have the following properties:

10-a) No restriction on specific types of probability distributions is imposed.

10-b) Independence of the knowledge of: 1) The input disturbance covariance C_w, 2) the input space-varying parameter $\beta(x)$, and 3) the output gain h.

10-c) Suitability for on-line identification.

10-d) Under no noise condition (i.e., d=0) the algorithm (SA-1) becomes independent of the knowledge of the parameters

$\{\gamma_m(x)\ ;\ m=0,\ldots,N\}$ and a linear combination of them can also be identified, as remarked on pages 115 and 121.

11) CONVERGENCE SPEED: The examples confirmed that the identification procedure has a good convergence speed, which can be accelerated still further by changing the sequences $\lambda(n)$ and $\mu(n)$ (i.e., by choosing optimal sequences).

12) SIMILARITIES WITH OTHER METHODS: Although this seems to be the first attempt to identify distributed systems in a stochastic environment (random inputs and noisy observations) without imposing restrictions on probability distributions, some common points with previous works can be pointed out (for details see section 2.3):

12-a) A slightly similar deterministic version of the parameter explicit technique, that uses the DPS (with constant parameters) reduced to a set of algebraic equations, was applied by Collins and Khatri [2].

12-b) Stochastic approximation algorithms, as a searching scheme for finding estimates previously obtained by minimizing a performance criterion, were used by Zhivoglyadov and Kaipov [3]-[6] and Carpenter, Wozny and Goodson [7]. They considered noisy observations, but not random inputs.

12-c) In [8], [9] Tzafestas considered random inputs but assumed perfect observation of the state.

12-d) Balakrishnan [10] presented a rigorous theoretical formulation for a particular DPS identification problem in a Gaussian stochastic environment.

Suggestions for further research: Some areas in the DPS identification field where further work seems needed have already been commented on section 2.4. Here we extend those remarks, regarding mainly the method developed in chapter 4, by suggesting the following topics for further research:

1) Other techniques for reduction to a finite-dimensional state space (stage I), such as finite element methods or even more elaborated finite-differences, could be investigated towards the applicability of our stochastic approximation approach in stage II.

2) As commented before, the identification procedure can be accelerated by changing the sequences $\lambda(n)$ and $\mu(n)$ appearing in (SA-1) and (SA-2). In this way, optimization studies could be done in order to determine a pair of sequences $(\lambda^*(n), \mu^*(n))$, among those satisfying the conditions required in (T - 4.1), which maximizes the algorithm convergence speed.

3) More research is also needed regarding optimal placement of a fixed number of observation points.

4) In proposition (P - 4.1) we presented a basic observation equation. A similar relation, also involving the state, could be formulated as follows:

$$z_k(\tau_n) = h\langle \underline{a}_k \, ; \, \underline{u}_k(n)\rangle + \nu_k(\tau_n)$$

where

$$\underline{u}_k(n) = (u_{k-\bar{M}}(n), \ldots, u_{k+\bar{M}}(n)) \; : \quad \text{a random vector in } R^{M+1}.$$

In this way we get similar algorithms as in (SA-1) and (SA-2), with $\underline{z}_k(n+N_t)$ replaced by $\underline{u}_k(n+N_t)$ and $\underline{\varepsilon}_k = \underline{0}$. This approach presents three

Suggestions for further research: Some areas in the DPS identification field where further work seems needed have already been commented on section 2.4. Here we extend those remarks, regarding mainly the method developed in chapter 4, by suggesting the following topics for further research:

1) Other techniques for reduction to a finite-dimensional state space (stage I), such as finite element methods or even more elaborated finite-differences, could be investigated towards the applicability of our stochastic approximation approach in stage II.

2) As commented before, the identification procedure can be accelerated by changing the sequences $\lambda(n)$ and $\mu(n)$ appearing in (SA-1) and (SA-2). In this way, optimization studies could be done in order to determine a pair of sequences $(\lambda^*(n), \mu^*(n))$, among those satisfying the conditions required in (T - 4.1), which maximizes the algorithm convergence speed.

3) More research is also needed regarding optimal placement of a fixed number of observation points.

4) In proposition (P - 4.1) we presented a basic observation equation. A similar relation, also involving the state, could be formulated as follows:

$$z_k(\zeta_n) = h\langle \underline{a}_k ; \underline{u}_k(n)\rangle + \nu_k(\zeta_n)$$

where

$$\underline{u}_k(n) = (u_{k-\tilde{M}}(n), \ldots, u_{k+\tilde{M}}(n)) \; : \quad \text{a random vector in } R^{M+1}.$$

In this way we get similar algorithms as in (SA-1) and (SA-2), with $\underline{z}_k(n+N_t)$ replaced by $\underline{u}_k(n+N_t)$ and $\underline{\varepsilon}_k = \underline{0}$. This approach presents three basic advantages:

1. The "a priori" knowledge of both noise statistics and the parameters $\{\gamma_m(x)\ ;\ m=0,\dots,N\}$ would no longer be necessary.

2. Space-varying observation noise (i.e, $v_k(n)$) could be considered.

3. The "indispensable" (in the former approach) information supplied by finite-difference techniques (i.e., $\sum_{m=1}^{M+1} a_k^m$ constant. See equation (11) in chapter 4 and also remark 3 on page 115) would not be required in this case, and so a wider class of approximation techniques could be applied in stage I.

On the other hand, the accessibility of the state $\underline{u}_k(n)$ must be assumed, which represents the main disadvantage. In order to by-pass the state accessibility requirement, the identification could be done by using in the new version of (SA-1) and (SA-2) estimates $\hat{\underline{u}}_k(n)$, instead of $\underline{u}_k(n)$, obtained by means of the Kalman-Bucy Filter. This approach has been previously considered for LPS identification in [11].

5) It would be useful to have some comparison of effectiveness of the different approaches for the DPS identification problem. In performing such comparisons, one must keep in mind that: The literature in this field considers a wide range of particular models operating in quite diverse conditions. So a critical evaluation involving a large number of methods could become a difficult task.

REFERENCES

CHAPTER 1

[1] ZADEH L. A. - "From Circuit Theory to System Theory", Proc. IRE, Vol. 50, pp. 856-865, May 1962.

[2] BELLMAN R. and ÅSTRÖM K. J. - "On Structural Identifiability", Math. Biosci., Vol. 1, pp. 329-339, 1969.

[3] ATHANS M. and FALB P. L. - "Optimal Control", McGraw-Hill, 1966.

[4] KALMAN R. E., FALB P. L. and ARBIB M. A. - "Topics in Mathematical System Theory", McGraw-Hill, 1969.

[5] HIRSCH M. W. and SMALE S. - "Differential Equations, Dynamical Systems, and Linear Algebra", Academic Press, 1974.

[6] MESAROVIC M. D. and TAKAHAR Y. - "General System Theory: Mathematical Foundations", Academic Press, 1975.

[7] GRENADER U. and ROSEMBLATT M. - "Statistical Analysis of Stationary Time Series", Wiley, 1957.

[8] WHITTLE P. - "Prediction and Regulation", Van Nostrand, 1963.

[9] JENKINS G. and WATTS D. - "Spectral Analysis and Its Applications", Holden-Day, 1968.

[10] WANG P. K. C. - "Control of Distributed Parameter Systems", in Advances in Control Systems, Vol. 1, (Ed. by C. T. Leondes), Academic Press, 1964.

[11] SCHWARZ R. J. and FIEDLAND B. - "Linear Systems", McGraw-Hill, 1965.

[12] COOPER G. R. and MC GILLEN C. D. - "Methods of Signals and System Analysis", Holt, R. & W., 1967.

[13] LEON B. J. - "Lumped Systems", Holt, R. & W., 1968.

[14] BUTKOVSKIY A. G. - "Distributed Control Systems", Elsevier, 1969.

[15] MC GILLEN C. D. and COOPER G. R. - "Continuous and Discrete Signal and Systems Analysis", Holt, R. & W., 1974.

[16] CHEN C. T. - "Analysis and Systhesis of Linear Control Systems", Holt, R. & W., 1975.

[17] ROBINSON A. C. - "A Survey of Optimal Control of Distributed Parameter Systems", Automatica, Vol. 7, pp. 371-389, May 1971.

[18] DERUSSO P. M., ROY B. J. and CLOSE C. M. - "State Variables for Engineers", Wiley, 1965.

[19] OGATA K. - "State Space Analysis of Control Systems", Prentice-Hall, 1967.

[20] BROCKETT R. W. - "Finite Dimensional Linear Systems", Wiley, 1970.

[21] CHEN C. T. - "Introduction to Linear System Theory", Holt, R. & W., 1970.

[22] KIRK D. E. - "Optimal Control Theory: An Introduction", Prentice-Hall, 1970.

[23] ROSENBROCK H. H. - "State Space and Multivariable Theory", Nelson, 1970.

[24] WIBERG D. M. - "State Space and Linear Systems", McGraw-Hill, 1970.

[25] DIRECTOR S. W. and ROHRER R. A. - "Introduction to System Theory", McGraw-Hill, 1972.

[26] WOLOVICH W. A. - "Linear Multivariate Systems", Springer-Verlag, 1974.

[27] ZADEH L. A. and DESOER C. A. - "Linear System Theory: The State Space Approach", McGraw-Hill, 1963.

[28] ZADEH L. A. - "The Concept of State in System Theory", in Views on General System Theory, (Ed. by M. Mesarovic), Wiley, 1964.

[29] BALAKRISHNAN A. V. - "Foundations of the State Space Theory of Continuous System", J. of Comp. & Syst. Sci., Vol. 1, No. 1, pp. 91-116, Mar. 1967.

[30] BALAKRISHNAN A. V. - "State Space Theory of Linear Time-Varying Systems", in System Theory, (Ed. by L. A. Zadeh and E. Polak), McGraw-Hill, 1969.

[31] ZADEH L. A. - "The Concept of System, Aggregate, and State in System Theory", in System Theory, (Ed. by L. A. Zadeh and E. Polak), McGraw-Hill, 1969.

[32] KALMAN R. E. - "A New Approach to Linear Filtering and Prediction Problems", Trans. ASME, J. of Basic Eng., Vol. 82-D, No. 1, pp. 35-45, Mar. 1960.

[33] KALMAN R. E. and BUCY R. S. - "New Results in Linear Filtering and Prediction Theory", Trans. ASME, J. of Basic Eng., Vol. 83-D, No. 1, pp. 95-108, Mar. 1961.

[34] TOU J. T. - "Optimum Design of Digital Control Systems", Academic Press, 1963.

[35] LEE R. C. K. - "Optimal Estimation, Identification and Control", MIT Press, 1964.

[36] DEUTSCH R. - "Estimation Theory", Prentice-Hall, 1965.

[37] SORENSON H. W. - "Kalman Filtering Techniques", in Advances in Control Systems, Vol. 3, (Ed. by C. T. Leondes), Academic Press, 1966.

[38] AOKI M. - "Optimization of Stochastic Systems", Academic Press, 1967.

[39] LIEBELT P. B. - "An Introduction to Optimal Estimation", Addison-Wesley, 1967.

[40] BUCY R. S. and JOSEPH P. D. - "Filtering for Stochastic Process with Application to Guidance", Interscience, 1968.

[41] LUENBERGER D. G. - "Optimization by Vector Space Methods", Wiley, 1968.

[42] SAGE A. P. - "Optimum Systems Control", Prentice-Hall, 1968.

[43] VAN TREES H. L. - "Detection, Estimation and Modulation Theory: Part I", Wiley, 1968.

[44] BRYSON A. E., Jr. and HO H. C. - "Applied Optimal Control", Ginn, 1969.

[45] DEUTSCH R. - "Systems Analysis Techniques", Prentice-Hall, 1969.

[46] MC CAUSLAND I. - "Introduction to Optimal Control", Wiley, 1969.

[47] MEDITCH J. S. - "Stochastic Optimal Linear Estimation and Control", McGraw-Hill, 1969.

[48] MORRISON N. - "Introduction to Sequential Smoothing and Prediction", McGraw-Hill, 1969.

[49] NAHI N. E. - "Estimation Theory and Applications", Wiley, 1969.

[50] ÅSTRÖM K. J. - "Introduction to Stochastic Control Theory", Academic Press, 1970.

[51] JAZWINSKI A. H. - "Stochastic Processes and Filtering Theory", Academic Press, 1970.

[52] KUO B. C. - "Discrete-Data Control Systems", Prentice-Hall, 1970.

[53] SPEED C. B., BROWN R. F. and GOODWIN G. C. - "Control Theory: Identification and Optimal Control", Oliver & Boyd, 1970.

[54] TAKAHASHI Y., RABINS M. J. and AUSLANDER D. M. - "Control and Dynamic Systems", Addison-Wesley, 1970.

[55] ANDERSON B. D. O. and MOORE J. B. - "Linear Optimal Control", Prentice-Hall, 1971.

[56] KUSHNER H. - "Introduction to Stochastic Control", Holt, R. & W., 1971.

[57] LEWIS T. D. and ODELL P. L. - "Estimation in Linear Models", Prentice-Hall, 1971.

[58] SAGE A. P. and MELSA J. L. - "Estimation Theory with Application to Communication and Control", McGraw-Hill, 1971.

[59] WONG E. - "Stochastic Processes in Information and Dynamical Systems", McGraw-Hill, 1971.

[60] DESAI R. C. and LALWANI C. S. - "Identification Techniques", Tata McGraw-Hill, 1972.

[61] KWAKERNAAK H. and SIVAN R. - "Linear Optimal Control Systems", Wiley-Interscience, 1972.

[62] BALAKRISHNAN A. V. - "Stochastic Differential Systems I - Filtering and Control: A Function Space Approach", Lectures Notes in Economics and Mathematical Systems, Vol. 84, Springer-Verlag, 1973.

[63] MELSA J. L. and SAGE A. P. - "An Introduction to Probability and Stochastic Processes", Prentice-Hall, 1973.

[64] MENDEL J. M. - "Discrete Techniques of Parameter Estimation", Dekker, 1973.

[65] ARNOLD L. - "Stochastic Differential Equations: Theory and Application", Wiley, 1974.

[66] EYKHOFF P. - "System Identification: Parameter and State Estimation", Wiley, 1974.

[67] JACOBS O. L. R. - "Introduction to Control Theory", Claredon Press-Oxford, 1974.

[68] MC GARTY T. P. - "Stochastic Systems & State Estimation", Wiley, 1974.

[69] KAILATH T. - "A View of Three Decades of Linear Filtering Theory", IEEE Trans., Inf. Theory, Vol. IT-20, No. 2, pp. 146-181, Mar. 1974.

[70] BENSOUSSAN A. - "Sur L'Identification et le Filtrage de Systèmes Gouvernés par des Équations aux Dérivées Partielles", IRIA, Chaier No. 1, Fév. 1969.

[71] BENSOUSSAN A. - "Identification de Systèmes Gouvernés par des Équations aux Dérivées Partielles", in Computing Methods in Optimization Problems, Vol. 2, (Ed. by L. A. Zadeh, L. W. Neustadt and A. V. Balakrishnan), Academic Press, 1969.

[72] BENSOUSSAN A. - "Filtrage Optimal des Systèmes Lineaires", Dunod, 1971.

[73] PHILLIPSON G. A. - "Identification of Distributed Parameter Systems", Elsevier, 1971.

[74] CURTAIN R. F. - "A Survey of Infinite Dimensional Filtering", SIAM Review, Vol. 17, No. 3, pp. 395-411, Jul. 1975.

[75] MEDITCH J. S. - "On State Estimation for Distributed Parameter Systems", J. Franklin Inst., Vol. 290, pp. 49-59, Jul. 1970.

[76] MEDITCH J. S. - "Least Squares Filtering and Smoothing for Linear Distributed Parameter Systems", Automatica, Vol. 7, pp. 315-322, May 1971.

[77] BALAKRISHNAN A. V. - "Stochastic System Identification Techniques", in Stochastic Optimization and Control, (Ed. by H. F. Karreman), Wiley, 1968.

[78] BOX G. E. P. and JENKINS G. M. - "Time-Series Analysis: Forecasting and Control", Holden-Day, 1970.

[79] SAGE A. P. and MELSA J. L. - "System Identification", Academic Press, 1971.

[80] GRAUPE P. - "Identification of Systems", Van Nostrand,1972.

[81] EYKHOFF P. - "Some Fundamental Aspects of Process Parameter Estimation", IEEE Trans., Autom. Control, Vol. AC-8, pp. 347-357, Oct. 1963.

[82] EYKHOFF P., VAN DER GRINTEN M. E. M., KWAKERNAAK H. and VELTEMAN B. P. - "System Modeling and Identification", Survey Paper, Proc. 3rd IFAC Congress, London, Jun. 1966.

[83] EYKHOFF P. - "Process Parameter and State Estimation", Automatica, Vol. 4, pp. 205-233, May 1968.

[84] ÅSTRÖM K. J. - "Lectures on the Identification Problem - The Least-Squares Method", Lund Inst. of Tech., Div. of Autom. Control, Report No. 6806, Sep. 1968.

[85] BALAKRISHNAN A. V. and PETERKA V. - "Identification in Automatic Control Systems", Automatica, Vol. 5, pp. 817-829, Nov. 1969.

[86] BEKEY G. A. - "System Identification an Introduction and a Survey", Simulation, pp. 151-166, Oct. 1970.

[87] NIEMAN R. E., FISHER D. G., SEBORG D. E. - "A Review of Process Identification and Parameter Estimation", Int. J. of Control, Vol. 13, No. 2, pp. 209-264, Feb. 1971.

[88] ÅSTRÖM K. J. and EYKHOFF P. - "System Identification - A Survey", Automatica, Vol. 7, pp. 123-162, Mar. 1971.

[89] SAGE A. P. - "System Identification - History, Methodology, Future Prospects", in System Identification of Vibrating Structures, (Ed. by W. D. Pilkey and R. Cohen), ASME, 1972.

[90] RAJBMAN N. S. - "The Application of Identification Methods - Survey", 3rd IFAC Symp. on Ident. & Syst. Param. Est., The Hague/Delft, pp. 1-48, Jun. 1973 (also in Automatica, Vol. 12, Jan. 1976).

[91] ALBUQUERQUE J. P. A. - "Identification Methods", Chapter II of: Sensitivity to Structural Errors in Randon Processes Identification Performance, Ph.D. Thesis, MIT, Sep. 1973.

[92] ALLISON J. S. - "On the Comparison of Two Methods of Off-Line Parameter Identification", J. of Math. Anal. & Appl., Vol. 18, No. 2, pp. 229-237, May 1967.

[93] CUENOD M. and SAGE A. P. - "Comparison of Some Methods Used for Process Identification", IFAC Symp.; Ident. in Autom. Control Syst., Prague, Jun. 1967.

[94] BOWLES R. L. and STRAETER T. A. - "System Identification: Computational Considerations", in System Identification of Vibrating Structures, (Ed. by W. D. Pilkey and R. Cohen), ASME, 1972.

[95] ISERMANN R., BAUR U., BAMBERGER W., KNEPPO P. and SIEBERT H. - "Comparison of Six On-Line Identification Methods", Automatica, Vol. 10, pp. 81-103, Jan. 1974.

[96] SARIDIS G. N. - "Comparison of Six On-Line Identification Algorithms", Automatica, Vol. 10, pp. 69-79, Jan. 1974.

[97] MURTHY D. N. P. - "Comparison of Various Methods for Estimating the Parameters Characterizing Noise in Discrete-Time Dynamical Systems", Int. J. of Syst. Sci., Vol. 5, No. 2, pp. 101-115, Feb. 1974.

[98] SINHA N. K. and SEN A. - "Critical Evaluation of Online Identification Methods", Proc. IEE, Vol. 122, No. 10, pp. 1153-1158, Oct. 1975.

[99] GOODSON R. E. and POLIS M. P. - "Parameter Identification in Distributed Systems: A Synthesising Overview", in Identification of Parameters in Distributed Systems, (Ed. by R. E. Goodson and M. P. Polis), ASME, 1974 (also in Proc. IEEE, Jan. 1976).

CHAPTER 2

[1] WANG P. K. C. - "Control of Distributed Parameter Systems", in Advances in Control Systems, Vol. 1, (Ed. by C. T. Leondes), Academic Press, 1964.

[2] BUTKOVSKIY A. G. - "Distributed Control Systems", Elsevier, 1969.

[3] GOODSON R. E. and POLIS M. P. - "Parameter Identification in Distributed Systems: A Synthesising Overview", in Identification of Parameters in Distributed Systems, (Ed. by R. E. Goodson and M. P. Polis), ASME, 1974.

[4] FORSYTHE G. E. and WASOW W. R. - "Finite-Difference Methods for Partial Differential Equations", Wiley, 1960.

[5] DOUGLAS J., Jr. - "A Survey of Numerical Methods for Parabolic Differential Equations", in Advances in Computer, Vol. 2, (Ed. by F. L. ALT), Academic Press, 1961.

[6] FOX L., (Ed.) - "Numerical Solutions of Ordinary and Partial Differential Equations", Pergamon Press, 1962.

[7] MIKHLIN S. G. - "Variational Methods in Mathematical Physics", Pergamon Press, 1964.

[8] COLLATZ L. - "The Numerical Treatment of Differential Equations", (3rd. Ed.), Springer-Verlag, 1966.

[9] COLLATZ L. - "Functional Analysis and Numerical Mathematics", Academic Press, 1966.

[10] ISAACSON E. and KELLER H. B. - "Analysis of Numerical Methods", Wiley, 1966.

[11] MIKHLIN S. G. and SMOLITSKIY K. L. - "Approximate Methods for Solutions of Differential and Integral Equations", Elsevier, 1967.

[12] AMES W. F. - "Numerical Methods for Partial Differential Equations", Nelson, 1969.

[13] VON ROSENBERG D. U. - "Methods for Numerical Solution of Partial Differential Equations", Elsevier, 1969.

[14] BLUM E. K. - "Numerical Analysis and Computation: Theory and Practice", Addisson Wesley, 1972.

[15] DESAI R. C. and LALWANI C. S. - "Distributed Parameter Estimation", Chapter 7 of Identification Techniques, Tata McGraw-Hill, 1972.

[16] STOLZ G., Jr. - "Numerical Solutions to an Inverse Problem of Heat Conduction for Simple Shapes", Trans. ASME, J. of Heat Transfer, Vol. 82-C, No. 1, pp. 20-26, Feb. 1960.

[17] FRANK J. - "An Application of Least-Squares Method to the Solution of the Inverse Problem of Heat Conduction", Trans. ASME, J. of Heat Transfer, Vol. 85-C, No. 4, pp. 378-379, Nov. 1963.

[18] BURGGRAF O. R. - "An Exact Solution of the Inverse Problem in Heat Conduction Theory and Applications", Trans. ASME, J. of Heat Transfer, Vol. 86-C, No. 3, pp. 373-382, Aug. 1964.

[19] SPARROW E. M., HAJI-SHEIKH A. and LUNDGREN T. S. - "The Inverse Problem in Transient Heat Conduction", Trans. ASME, J. of Appl. Mech., Vol. 31-E, No. 3, pp. 369-375, Sep. 1964.

[20] BECK J. V. - "The Optimum Analytical Design of Transient Experiments for Simultaneous Determinations of Thermal Conductivity and Specific Heat", Ph.D. Thesis, Dept. of Mech. Eng., Michigan State University, 1964.

[21] BECK J. V. - "Calculation of Thermal Diffusivity from Temperature Measurements", Trans. ASME, J. of Heat Transfer, Vol. 85-C, No. 2, pp. 181-182, May 1963.

[22] BECK J. V. - "Transient Sensivity Coefficient for the Thermal Contact Conductance", Int. J. of Heat and Mass Transfer, Vol. 10, pp. 1615-1617, 1967.

[23] BECK J. V. - "Surface Heat Flux Determination Using an Integral Method", Nucl. Eng. Design, Vol. 1, pp. 170-178, 1968.

[24] BECK J. V. - "Analytical Determination of Optimum Transient Experiments for the Measurement of Thermal Properties", Proc. 3rd Int. Heat Transfer Conf., Vol. IV, pp. 301-329, 1966.

[25] BECK J. V. - "Nonlinear Estimation Applied to the Nonlinear Inverse Heat Conduction Problem", Int. J. of Heat and Mass Transfer, Vol. 13, pp. 703-716, Apr. 1970.

[26] BECK J. V. - "Sensitivity Coefficients Utilized in Nonlinear Estimation with Small Parameters in a Heat Transfer Problem", Trans. ASME, J. of Basic Eng., Vol. 92-D, No. 2, pp. 215-222, Jun. 1970.

[27] ANGEL E. - "Inverse Boundary-Value Problems: Elliptic Equations", J. of Math. Anal. & Appl., Vol. 30, No. 1, pp. 86-98, Apr. 1970.

[28] JONES B. F., Jr. - "The Determination of a Coefficient in a Parabolic Differential Equation", Ph.D. Thesis, Rice University, 1961.

[29] JONES B. F., Jr. - "The Determination of a Coefficient in a Parabolic Differential Equation. Part I - Existence and Uniqueness", J. of Math. & Mech., Vol. 11, No. 6, pp. 907-918, 1962.

[30] DOUGLAS J., Jr. and JONES B. F., Jr. - "The Determination of a Coefficient in a Parabolic Differential Equation. Part II - Numerical Approximations", J. of Math. & Mech., Vol. 11, No. 6, pp. 919-926, 1962.

[31] JONES B. F., Jr. - "Various Methods for Finding Unknown Coefficients in Parabolic Differential Equations", Comm. on Pure & Appl. Math., Vol. 16, No. 1, pp. 33-44, Feb. 1963.

[32] CANNON J. R., DOUGLAS J., Jr. and JONES B. F., Jr. - "Determination of the Diffusitivity of an Isotropic Medium", Int. J. of Eng. Sci., Vol. 1, No. 4, pp. 453-455, 1963.

[33] CANNON J. R. and JONES B. F., Jr. - "Determination of the Diffusivity of an Anisotropic Medium", Int. J. of Eng. Sci., Vol. 1, No. 4, pp. 457-460, 1963.

[34] CANNON J. R. - "Determination of Unknown Coefficients in a Parabolic Differential Equation", Duke Math. J., Vol. 30, pp. 313-324, 1963.

[35] CANNON J. R. - "Determination of Certain Parameters in Heat Conduction Problems", J. of Math. Anal. & Appl., Vol. 8, No. 2, pp. 188-201, Apr. 1964.

[36] CANNON J. R. - "Determination of Unknown Coefficients k(u) in the Equation $\nabla \cdot k(u) \nabla (u) = 0$ from Overspecified Boundary Data", J. of Math. Anal. & Appl., Vol. 18, No. 1, pp. 112-114, Apr. 1967.

[37] BELLMAN R., DETCHMENDY D., KAGIWADA H. and KALAB R. - "On the Identification of Systems and the Unscrambling of Data - III: One Dimensional Wave and Diffusion Processes", J. of Math. Anal. & Appl., Vol. 23, No. 1, pp. 173-182, Jul. 1968.

[38] CLEMENTS W. C., Jr. - "A Note on the Determination of the Parameters of the Longitudinal Dispersion Model from Experimental Data", Chem. Eng. Sci., Vol. 24, pp. 957-963, 1969.

[39] DREIFKE G. E. and HOUGEN J. O. - "Experimental Determination of Systems Dynamics by Pulse Methods", J.A.C.C., pp. 452-458, Jun. 1963.

[40] SANATHANAN C. K. - "On Synthesis of Space Dependent Transfer Function", IEEE Trans., Autom. Control, Vol. AC-11, No. 4, pp. 724-729, Oct. 1966.

[41] WILLIAMS J. A., ADLER R. J. and ZOLNER W.J. - "Parameter Estimation of Unsteady State Distributed Models in the Laplace Domain", Ind. & Eng. Chem. Fundamentals, Vol. 9, No. 2, pp. 193-197, 1970.

[42] JUMARIE G. - "Identification on Nonlinear Distributed Systems by Using an Inverse Describing Function Method", 3rd IFAC Symp. on Ident. & Syst. Param. Est., The Hague/Delft, pp. 649-660, Jun. 1973.

[43] KOZHINSKY O. S. and RAJBMAN N. S. - "Identification of Didtributed Plants", IFAC Symp. on the Control of D.P.S., Banff, p-13.5, Jun. 1971.

[44] RAJBMAN N. S. - "The Application of Identification Methods", 3rd IFAC Symp. on Ident. & Syst. Param. Est., The Hague/Delft, pp. 1-48, Jun. 1973 (also in Automatica, Vol. 12, pp. 73-95, Jan. 1976).

[45] PERDREAUVILLE F. J. - "Identification of Systems Described by Partial Differential Equations", Ph.D. Thesis, Purdue University Lafayette, Indiana, Jan. 1966.

[46] PERDREAUVILLE F. J. and GOODSON R. E. - "Identification of Systems Described by Partial Differential Equations", Trans. ASME, J. of Basic Eng., Vol. 88-D, No. 2, pp. 463-468, Jun. 1966.

[47] SHINBROT M. - "On Analysis of Linear and Nonlinear Dynamical Systems from Transient-Response Data", NACA Tech. Notes, TN 3288, 1954.

[48] LOEB J. M. and CAHEN G. M. - "More About Process Identification", IEEE Trans., Autom. Control, Vol. 10, No. 4, pp. 359-361, Jul. 1965.

[49] TAKAYA K. - "The Use of Hermite Function for Systems Identification", IEEE Trans., Autom. Control, Vol. AC-13, No. 4, pp. 446-447, Aug. 1968.

[50] ZHIVOGLYADOV V. P. and KAIPOV V. K. - "Application of the Method of Stochastic Approximations in the Problem of Identification", Autom. & Rem. Control, Vol. 27, No. 10, pp. 1702-1706, Oct. 1966.

[51] ZHIVOGLYADOV V. P. and KAIPOV V. K. - "Identification of Distributed Plants in the Presence of Noises", IFAC Symp. on Ident. of Autom. Control Syst., Prague, p-3.5, Jun. 1967.

[52] ZHIVOGLYADOV V. P. and KAIPOV V. K. - "Accuracy of Distributed Systems Identification Algorithms", IFAC Symp. on Ident. & Proc. Param. Est., Prague, Jun. 1970.

[53] ZHIVOGLYADOV V. P., KAIPOV V. K. and TSIKUNOVA J. M. - "Stochastic Algorithms of Identification and Adaptive Control of Distributed Parameter Systems", IFAC Symp. on the Control of D.P.S., Banff, p-13.1, Jun. 1971.

[54] COLLINS P. L., Jr. - "Identification of Distributed Parameter Systems Using Finite-Differences", Ph.D. Dissertation, Montana State University, Bozeman, Mar. 1968.

[55] COLLINS P. L. and KHATRI H. C. - "Identification of Distributed Parameter Systems Using Finite-Differences", Trans. ASME, J. of Basic Eng., Vol. 91-D, No. 2, pp. 239-245, Jun. 1969.

[56] SEINFELD J. H. - "Identification of Parameters in Partial Differential Equations", Chem. Eng. Sci., Vol. 24, No. 1, pp. 65-74, Jan. 1969.

[57] SEINFELD J. H. and CHEN W. H. - "Estimation of Parameters in Distributed Systems from Noisy Experimental Data", Chem. Eng. Sci., Vol. 26, pp. 753-766, Jun. 1971.

[58] CHEN W. H. and SEINFELD J. H. - "Estimation of Spatially Varying Parameters in Partial Differential Equations", Int. J. of Control, Vol. 15, No. 3, pp. 487-495, Mar. 1972.

[59] CHEN W. H., GAVALAS G. R. and SEINFELD J. H. - "A New Algorithm for Automatic History Matching", 48th Ann. Meet. Soc. of Petroleum Eng., Las Vegas, Oct. 1973.

[60] SEINFELD J. H. and CHEN W. H. - "Estimation of Parameters in Distributed Systems", in Identification of Parameters in Distributed Systems, (Ed. by R. E. Goodson and M. P. Polis), ASME, 1974.

[61] CHEN W. H. - "Estimation of Parameters in Partial Differential Equations - With Application to Pretoleum Reservoir Description", Ph.D. Thesis, California Institute of Technology, 1974.

[62] GOODSON R. E. and KLEIN R. E. - "A Definition and Some Results for Distributed Systems Observability", IEEE Trans., Autom. Control, Vol. AC-15, No. 2, pp. 165-175, Apr. 1970.

[63] YU T. K. and SEINFELD J. H. - "Observability of a Class of Hyperbolic Distributed Parameter Systems", IEEE Trans., Autom. Control, Vol. AC-16, No. 5, pp. 495-496, Oct. 1971.

[64] SAKAWA Y. - "Observability and Related Problems for Partial Differential Equations of Parabolic Type", SIAM J. of Control, Vol. 13, No. 1, pp. 14-27, Jan. 1975.

[65] TZAFESTAS S. G. - "Identification of Stochastic Distributed Parameter Systems", Int. J. of Control, Vol. 11, No. 4, pp. 619-624, 1970.

[66] FAIRMAN F. W. and SHEN D. W. C. - "Parameter Identification for a Class of Distributed Systems", Int. J. of Control, Vol. 11, No. 6, pp. 929-940, 1970.

[54] COLLINS P. L., Jr. - "Identification of Distributed Parameter Systems Using Finite-Differences", Ph.D. Dissertation, Montana State University, Bozeman, Mar. 1968.

[55] COLLINS P. L. and KHATRI H. C. - "Identification of Distributed Parameter Systems Using Finite-Differences", Trans. ASME, J. of Basic Eng., Vol. 91-D, No. 2, pp. 239-245, Jun. 1969.

[56] SEINFELD J. H. - "Identification of Parameters in Partial Differential Equations", Chem. Eng. Sci., Vol. 24, No. 1, pp. 65-74, Jan. 1969.

[57] SEINFELD J. H. and CHEN W. H. - "Estimation of Parameters in Distributed Systems from Noisy Experimental Data", Chem. Eng. Sci., Vol. 26, pp. 753-766, Jun. 1971.

[58] CHEN W. H. and SEINFELD J. H. - "Estimation of Spatially Varying Parameters in Partial Differential Equations", Int. J. of Control, Vol. 15, No. 3, pp. 487-495, Mar. 1972.

[59] CHEN W. H., GAVALAS G. R. and SEINFELD J. H. - "A New Algorithm for Automatic History Matching", 48th Ann. Meet. Soo. of Petroleum Eng., Las Vegas, Oct. 1973.

[60] SEINFELD J. H. and CHEN W. H. - "Estimation of Parameters in Distributed Systems", in Identification of Parameters in Distributed Systems, (Ed. by R. E. Goodson and M. P. Polis), ASME, 1974.

[61] CHEN W. H. - "Estimation of Parameters in Partial Differential Equations - With Application to Pretoleum Reservoir Description", Ph.D. Thesis, California Institute of Technology, 1974.

[62] GOODSON R. E. and KLEIN R. E. - "A Definition and Some Results for Distributed Systems Observability", IEEE Trans., Autom. Control, Vol. AC-15, No. 2, pp. 165-175, Apr. 1970.

[63] YU T. K. and SEINFELD J. H. - "Observability of a Class of Hyperbolic Distributed Parameter Systems", IEEE Trans., Autom. Control, Vol. AC-16, No. 5, pp. 495-496, Oct. 1971.

[64] SAKAWA Y. - "Observability and Related Problems for Partial Differential Equations of Parabolic Type", SIAM J. of Control, Vol. 13, No. 1, pp. 14-27, Jan. 1975.

[65] TZAFESTAS S. G. - "Identification of Stochastic Distributed Parameter Systems", Int. J. of Control, Vol. 11, No. 4, pp. 619-624, 1970.

[66] FAIRMAN F. W. and SHEN D. W. C. - "Parameter Identification for a Class of Distributed Systems", Int. J. of Control, Vol. 11, No. 6, pp. 929-940, 1970.

[67] KANTOROVICH L. V. and KRYLOV V. J. - "Approximate Methods of Higher Analysis", Interscience, 1964.

[68] LESER T. - "Recent Sovier Contributions to Mathematics", Macmillan, 1962.

[69] HICKS J. S. and WEI J. - "Numerical Solution of Parabolic Partial Differential Equations with Two Point Boundary by Use of the Method of Lines", J. of the Assoc. for Comp. Mach., Vol. 14, No. 3, pp. 549-562, Jul. 1967.

[70] CARPENTER W. T. - "The Identification of Distributed Parameter Systems", Ph.D. Thesis, Purdue University, Lafayette, Aug. 1969.

[71] WOZNY M. J. and CARPENTER W. T. - "A Characteristic Approach to Parameter Identification in Distributed Systems", IEEE Trans., Autom. Control, Vol. AC-14, No. 4, pp. 422, Aug. 1969.

[72] CARPENTER W. T., WOZNY M. J. and GOODSON R. E. - "Distributed Parameter Identification Using the Method of Characteristics", Trans. ASME, J. of Dyn. Syst., Meas. & Control, Vol. 93-G, No. 2, pp. 73-78, Jun. 1971.

[73] COURANT R. and HILBERT D. - "Methods of Mathematical Physics", Vol. II, Interscience, 1962.

[74] GARABIDEAN P. R. - "Partial Differential Equations", Wiley, 1964.

[75] RUBAN A. I. - "Identification of Distributed Dynamic Objects on the Basic of a Sensitivity Algorithm", Eng. Cybernetics, Vol. 9, No. 6, pp. 1137-1142, Nov.- Dec. 1971.

[76] POLIS M. P. - "On Problems of Parameter Identification for Distributed Parameter Systems with Results Using Galerkin's Criterion", Ph.D. Thesis, Purdue University, Lafayette, 1972.

[77] POLIS M. P., WOZNY M. J. and GOODSON R. E. - "Identification of Parameters in Distributed Systems Using Galerkin's Method", IFAC Symp. on the Control of DPS, Banff, p-13.2, Jun. 1971.

[78] POLIS M. P., GOODSON R. E. and WOZNY M. J. - "Parameter Identification for the Bean Equation Using Galerkin's Criterion", IEEE Dec. & Control Conf., pp. 391-395, Dec. 1972.

[79] POLIS M. P., GOODSON R. E. and WOZNY M. J. - "On Parameter Identification for Distributed Systems Using Galerkin's Criterion", Automatica, Vol. 9, No. 1, pp. 53-64, Jan. 1973.

[80] DI PILLO G. and GRIPPO L. - "Application of the Epsilon Technique to the Identification of Distributed-Parameter Systems", J. of Opt. Tech. & Appl., Vol. 11, No. 1, pp. 84-99, 1973.

[81] DI PILLO G., GRIPPO L. and LUCERTINI M. - "Identification of Distributed Parameter Systems Using the Epsilon Method: A New Algorithm", 3rd IFAC Symp. on Ident. & Syst. Param. Est., The Hague/Delft, pp. 687-690, Jun. 1973.

[82] BALAKRISHNAN A. V. - "A New Computing Technique in Systems Identification", J. of Comp. & Syst. Sci., Vol. 2, No. 1, pp. 102-116, Jun. 1968.

[83] HAMZA M. H. and SHEIRAH M. A. - "On-Line Identification of Distributed Parameter Systems", Automatica, Vol. 9, No. 6, pp. 689-698, Nov. 1973 (also in IFAC Symp. on the Control of D.P.S., Banff, p-13.6, Jun. 1971).

[84] BHAGAVAN B. K. and NARDIZZI L. R. - "Parameter Identification in Linear Distributed Systems", IEEE Trans., Autom. Control, Vol. AC-18, No. 6, pp. 677-679, Dec. 1973.

[85] KUBRUSLY C. S. and CURTAIN R. F. - "Identification of Noisy Distributed Parameter Systems Using Stochastic Approximation", Control Theory Centre, Univ. of Warwick, CTC Report No. 37, May 1975, (to appear in Int. J. of Control).

[86] DIAMESSIS S. E. - "On the Simultaneous Identification of Distributed Systems Parameters and of Boundary Conditions in the Form of Power Series and Chebychev Polynomials. Application to the Wave Equation of Eletromagnetics and Transmission Lines Theory", IFAC Symp. on the Control of D.P.S., Banff, p-13.4, Jun. 1971.

[87] LUCKINBILL D. L. and CHILDS B. - "Parameter Identification by Newton-Raphson Expansion of Partial Differential Equations", IFAC Symp. on the Control of D.P.S., Banff, p-13.3, Jun. 1971.

[88] SHERRY H. and SHEN D. W. C. - "Combined State and Parameter Estimation for Distributed Parameter Systems Using Discrete Observations", 3rd IFAC Symp. on Ident. & Syst. Param. Est., The Hague/Delft, pp. 671-677, Jun. 1973.

[89] SHRIDHAR M. and BALATONI N. A. - "Application of Cubic Splines to Systems Identification", 3rd IFAC Symp. on Ident. & Syst. Param. Est., The Hague/Delft, pp. 787-791, Jun. 1973.

[90] CHAUDHURI S. P. - "A Class of Distributed Parameter System Identification via Differential Approximation and Quasi Linearisation", 5th Hawaii Inter. Conf. on Syst. Sci., Univ. of Hawaii, Jan. 1973.

[91] CHAVENT G. - "Analyse Fonctionelle et Identification de Coefficients Répartis dans les Équation aux Dérivées Partielles", Thèse Doctorat d'Etat ès Sciences, Paris, 1971.

[92] CHAVENT G. - "Sur une Méthode de Resolution du Problème Inverse dans les Équation aux Dérivées Partielles Parabolique", Note CRAS Paris, T-260, Dec. 1969.

[93] CHAVENT G. - "Deux Résultats sur le Problème Inverse dans les Équation aux Dérivées Partielles du 2ème Ordre en t et sur l'Unicite de le Solution du Problème Inverse de la Diffusion", Note CRAS Paris, T-270, Jan. 1970.

[94] CHAVENT G. - "Une Méthode de Résolution de Problème Inverse dans les Équations aux Dérivées Partielles", Bull. de l'Acad. Polonaise des Sci., Vol. 18, No. 3, pp. 99-105, 1970.

[95] CHAVENT G. and CILLIGOT-TRAVAIN G. - "L'Identification des Systèmes", IRIA Bull., Nov. 1970.

[96] CILLIGOT-TRAVAIN G. - "Estimation Bayesienne et Identification des Systèmes Gouvernés par des Équation aux Dérivées Partielles", Thèse Docteur Ingenieur, Paris, 1970.

[97] CHAVENT G. - "Identification of Distributed Parameters", 3rd IFAC Symp. on Ident. & Syst. Param. Est., The Hague/Delft, pp. 649-660, Jun. 1973.

[98] CHAVENT G., DUPUY M. and LEMONNIER P. - "History Matching by Use of Optimal Control Theory", 48th Ann. Meet. Soc. of Petroleum Eng., Las Vegas, Oct. 1973.

[99] CHAVENT G. and LEMONNIER P. - "Identification de la Non-Linearité d'une Équation Parabolique Quasi-Linéaire", Appl. Math. & Optimization, Vol. 1, No. 2, pp. 121-162, Springer-Verlag, 1974.

[100] CHAVENT G. - "Identification of Functional Parameters in Partial Differential Equations", in Identification of Parameters in Distributed Systems, (Ed. by R. E. Goodson and M. P. Polis), ASME, 1974.

[101] CHAVENT G. - "Estimation de Paramèteres Distribués dans le Équation aux Dérivées Partielles", Comp. Meth. in Appl. Sci. & Eng. - Part 2, Lecture Notes in Computer Science, Vol. 11, Springer-Verlag, 1974.

[102] BALAKRISHNAN A. V. - "Identification of Systems Subject to Random State Disturbances", 5th Conf. on Opt. Tech. - Part I, Lecture Notes in Computer Science, Vol. 3, Springer-Verlag, 1973.

[103] BALAKRISHNAN A. V. - "Identification-Inverse Problems for Partial Differential Equations: A Stochastic Formulation", 6th IFIP Conf. on Opt. Tech., Lecture Notes in Computer Science, Vol. 27, Springer-Verlag, 1975.

[104] BALAKRISHNAN A. V. - "Identification and Stochastic Control of a Class of Distributed Systems with Boundary Noise", IRIA Conf. on Stoch. Control and Comp. Modelling, Jun. 1974 (to be published by Springer-Verlag, 1975).

[105] LAVRENTIEV, ROMANOV, VASILIEV - "Multi-Dimensional Inverse Problems for Differential Equations", Lecture Notes in Mathematics, Vol. 167, Springer-Verlag, 1970.

[106] MARCHUK G. - "Perturbation Theory and the Statement of Inverse Problems", 5th Conf. on Opt. Tech. - Part II, Lecture Notes in Computer Science, Vol. 4, Springer-Verlag, 1973.

[107] LIONS J. L. - "Contrôle Optimal de Systèmes Gouvernés par des Équations aux Dérivées Partielles", Dunod, 1968.

[108] WARD E. D. - "Identifications of Parameters in Nonlinear Boundary Conditions of Distributed Systems with Linear Fields", Ph.D. Thesis, Purdue University, Lafayette, Aug. 1971.

[109] WARD E. D. and GOODSON R. E. - "Identification of Parameters in Nonlinear Boundary Conditions of Distributed Systems with Linear Fields", Trans. ASME, J. of Dyn. Syst., Meas. & Control, Vol. 95-G, No. 4, pp. 391-395, Dec. 1973.

[110] WOZNY M. J., CARPENTER W. T. and STEIN G. - "Identification of Green's Function for Distributed Parameter Systems", IEEE Trans., Autom. Control, Vol. AC-15, No. 1, pp. 155-157, Feb. 1970.

[111] CANNON J. R. - "Determination of an Unknown Heat Source from Overspecified Boundary Data", SIAM J. of Numer. Anal., Vol. 5, No. 2, pp. 275-286, Jun. 1968.

[112] SARIDIS G. N. and BADAVAS P. C. - "Identification of Distributed Parameter Systems by Stochastic Approximation", 7th Symp. on Adaptive Processes, University of California, Berkeley, Dec. 1968.

[113] SARIDIS G. N. and BADAVAS P. C. - "Identifying Solutions of Distributed Parameter Systems by Stochastic Approximation", IEEE Trans., Autom. Control, Vol. AC-15, No. 3, pp. 393-395, Jun. 1970.

[114] PHILLIPSON G. and MITTER S. K. - "State Identification of a Class of Linear Distributed Systems", 4th IFAC Congress, Warsaw, pp. 34-56, Jun. 1969.

[115] PHILLIPSON G. A. - "Identification of Distributed Parameter Systems", Elsevier, 1971.

[116] VEMURI V. and KARPLUS W. J. - "Identification of Non-Linear Parameters of Ground Water Basins by Hybrid Computation", Water Resources Research, Vol. 5, No. 1, Feb. 1969.

[117] DALE B. O. - "Multiparameter Identification and Optimization Methods for Linear Continuous Vibratory Systems", Ph.D Thesis, Purdue University, Lafayette, Jan. 1970.

[118] DALE B. O. and COHEN R. - "Multiparameter Identification in Linear Continuous Vibratory Systems", Trans. ASME, J. of Dyn. Syst., Meas. & Control, Vol. 93-G, No. 1, pp. 45-52, Mar. 1971.

[119] GILATH C. - "Determination of the Dilution of a Gas Discharged into a Ventilated Atmosphere", Int. J. of Appl. Rad. & Isotopes, Vol. 22, pp. 671-675, 1971.

[120] GILATH C. and SUTHL A. - "Concentration Dynamics in a Lake with a Water Current Flowing Through It", Proc. IAEA Symp. on the Use of Nucl. Tech. in the Meas. & Control of Env. Pollution, Vienna, pp. 483-496, 1971.

[121] GILATH C., et al. - "Radioisotope Tracer Techniques in the Investigation of Dispersion of Sewage and Disappearance Rate of Enteric Organisms in Coastal Waters", Proc. IAEA Symp. on the Use of Nucl. Tech. in the Meas. & Control of Env. Pollution, Vienna, pp. 673-689, 1971.

[122] KOIVO A. J. and PHILLIPS G. R. - "Identification of Mathematical Models for DO and BOD Concentrations in Polluted Streams from Noise Corrupted Measurements", Water Resources Research, Vol. 7, No. 4, pp. 853-862, Aug. 1971.

[123] KOIVO A. J. and PHILLIPS G. R. - "On Determination of BOD and Parameters in Polluted Stream Models from DO Measurements Only", Water Resources Research, Vol. 8, No. 2, Apr. 1972.

[124] PILKEY W. D. and COHEN R. (Ed.) - "System Identification of Vibrating Structures - Mathematical Models from Test Data", ASME, 1972.

[125] LEDEN B. - "Identification of Dynamics of a One-Dimensional Heat Diffusion Process", Lund Institute of Technology, Division of Auto-Control, Report 7121, Nov. 1971.

[126] LEDEN B., HAMZA M. H. and SHEIRAH M. A. - "Different Methods for Estimation of Thermal Diffusivity of a Heat Diffusion Process", 3rd IFAC Symp. on Ident. & Syst. Param. Est., The Hague/Delft, pp. 639-648, Jun. 1973.

[127] MALPANI S. N. and DONNELLY J. K. - "Identification of One Phase Flow Processes", Canadian J. of Chem. Eng., Vol. 50, No. 6, pp. 791-795, Dec. 1972.

[128] MALPANI S. N. and DONNELLY J. K. - "Identification of a Packed Absorption Column", Canadian J. of Chem. Eng., Vol. 51, No. 4, pp. 479-483, Aug. 1973.

[129] AGUILAR MARTIN J., ALENGRIN G. and HERNANDEZ C. - "Estimation of the Diffusion Coefficient of a Fluidised Bed Catalytic Reactor Using Non Linear Filtering Techniques", 3rd IFAC Symp. on Ident. & Syst. Param. Est., The Hague/Delft, pp. 255-264, Jun. 1973.

[130] AGUILAR MARTIN J., CONTINENTE G., ALENGRIN G. and HERNANDEZ C.- "Méthodes Recurrents d'Estimation de Modèles Dynamiques de Processes Chimiques Industriels", Fédération Européenne de Génie Chemique, Cong. Int., Paris, Avr. 1973.

[131] MC GREAVY C. and VAGO A. - "Estimation of Spatially Distributed Decay Parameters in a Chemical Reactor", 3rd IFAC Symp. on Ident. & Syst. Param. Est., The Hague/Delft, pp. 307-316, Jun. 1973.

[132] PINDER G. F. - "Functional Coefficients in the Analysis of Groundwater Flow", Water Resources Research, Vol. 9, No. 1, Jan. 1973.

[133] HUGHES R. D. - "Parameter Identification of Vibrating Second Order Hyperbolic Distributed Systems", Ph.D. Thesis, University of Illinois at Urbana, 1974.

[134] SPALDING G. R. - "A Green's Function Approach to the Identification of Linear DPS", Ph.D Thesis, Lehigh University, Bethlehem, 1974.

[135] COHEN R. - "Identification in Vibratory Systems: A Survey", in Identification of Parameters in Distributed Systems, (Ed. By R. E. Goodson and M. P. Polis), ASME, 1974.

[136] BENNETT R. J. - "Identification of the Minimal Representation of Linear Distributed Parameter Economic Systems", in Identification of Parameters in Distributed Systems, (Ed. by R. E. Goodson and M. P. Polis), ASME, 1974.

[137] STRANG G. and FIX G. J. - "An Analysis of the Finite Element Method", Prentice-Hall, 1973.

[138] ATHANS M. - "Towards a Practical Theory for Distributed Systems", IEEE Trans., Autom. Control, Vol. AC-15, No. 2, pp. 245-247, Apr. 1970.

*[139] ALVARADO F. L. and MUKUNDAN R. - "An Identification Problem in Distributed-Parameter Systems", Int. J. of System Sci., Vol. 1, No. 2, pp. 173-183, 1970.

*[140] TZAFESTAS S. G. - "Identification of Hybrid Distributed Parameter Systems", Int. J. of Control, No. 1, Vol. 13, pp. 145-154, Jan. 1971.

*[141] CANNON J. R. and DUCHATEAU P. - "Determining Unknown Coefficients in a Nonlinear Heat Conduction Problem", SIAM J. of Appl. Math., Vol. 24, No. 3, pp. 298-314, May 1973.

*[142] IKEDA S., MIYAMOTO S. and SAWARGI Y. - "Identification Method in Environmental Systems and its Applications to Water Polutions", Int. J. of System Sci., Vol. 5, No. 8, pp. 707-723, 1974.

*[143] GOODSON R. E. and POLIS M. P. - "A Survey of Parameter Identification in Distributed Systems", 6th IFAC Congress, Boston, p-8.2, Aug. 1975 (also in Proc. IEEE, Vol. 64, No. 1, pp. 43-61, Jan. 1976).

---

* These references came to the author's knowledge later when the chapter was fully typed, so they do not follow the previous order.

CHAPTER 3

[1] WONG E. - "Stochastic Processes in Information and Dynamical Systems", McGraw-Hill, 1971.

[2] ADAMS R. A. - "Sobolev Spaces", Academic Press, 1975.

[3] TREVES F. - "Basic Linear Partial Differential Equations", Academic Press, 1975.

[4] COURANT R. and HILBERT D. - "Methods of Mathematical Physics", Vol. II, Interscience, 1962.

[5] MIKHLIN S. G. - "Mathematical Physics: An Advanced Course", North-Holland, 1970.

[6] GARABEDIAN P. R. - "Partial Differential Equations", Wiley, 1964.

[7] FORSYTHE G. E. and WASOW W. R. - "Finite-Difference Methods for Partial Differential Equations", Wiley, 1960.

[8] SALVADORI M. G. and BARON M. L. - "Numerical Methods in Engineering", Prentice-Hall, 1961.

[9] COLLATZ L. - "The Numerical Treatment of Differential Equations", (3rd Ed.), Springer-Verlag, 1966.

[10] COLLATZ L. - "Functional Analysis and Numerical Mathematics", Academic Press, 1966.

[11] ISAACSON E. and KELLER H. B. - "Analysis of Numerical Methods", Wiley, 1966.

[12] WENDROFF B. - "Theoretical Numerical Analysis", Academic Press, 1966.

[13] HILDEBRAND F. B. - "Finite-Differences and Simulation", Prentice-Hall, 1968.

[14] AMES W. F. - "Numerical Methods for Partial Differential Equations", Nelson, 1969.

[15] MITCHELL A. R. - "Computational Methods in Partial Differential Equations", Wiley, 1969.

[16] SMITH G. D. - "Numerical Solution of Partial Differential Equations", Oxford Univ. Press, 1969.

[17] VON ROSENBERG D. U. - "Methods for Numerical Solution of Partial Differential Equations", Elsevier, 1969.

[18] GERALD C. F. - "Applied Numerical Analysis", Addison-Wesley, 1970.

CHAPTER 3

[1] WONG E. - "Stochastic Processes in Information and Dynamical Systems", McGraw-Hill, 1971.

[2] ADAMS R. A. - "Sobolev Spaces", Academic Press, 1975.

[3] TREVES F. - "Basic Linear Partial Differential Equations", Academic Press, 1975.

[4] COURANT R. and HILBERT D. - "Methods of Mathematical Physics", Vol. II, Interscience, 1962.

[5] MIKHLIN S. G. - "Mathematical Physics: An Advanced Course", North-Holland, 1970.

[6] GARABEDIAN P. R. - "Partial Differential Equations", Wiley, 1964.

[7] FORSYTHE G. E. and WASOW W. R. - "Finite-Difference Methods for Partial Differential Equations", Wiley, 1960.

[8] SALVADORI M. G. and BARON M. L. - "Numerical Methods in Engineering", Prentice-Hall, 1961.

[9] COLLATZ L. - "The Numerical Treatment of Differential Equations", (3rd Ed.), Springer-Verlag, 1966.

[10] COLLATZ L. - "Functional Analysis and Numerical Mathematics", Academic Press, 1966.

[11] ISAACSON E. and KELLER H. B. - "Analysis of Numerical Methods", Wiley, 1966.

[12] WENDROFF B. - "Theoretical Numerical Analysis", Academic Press, 1966.

[13] HILDEBRAND F. B. - "Finite-Differences and Simulation", Prentice-Hall, 1968.

[14] AMES W. F. - "Numerical Methods for Partial Differential Equations", Nelson, 1969.

[15] MITCHELL A. R. - "Computational Methods in Partial Differential Equations", Wiley, 1969.

[16] SMITH G. D. - "Numerical Solution of Partial Differential Equations", Oxford Univ. Press, 1969.

[17] VON ROSENBERG D. U. - "Methods for Numerical Solution of Partial Differential Equations", Elsevier, 1969.

[18] GERALD C. F. - "Applied Numerical Analysis", Addison-Wesley, 1970.

[19] YOUNG D. M. and GREGORY R. T. - "A Survey of Numerical Mathematics", Vol. II, Addison-Wesley, 1973.

[20] NAYLOR A. W. and SELL G. R. - "Linear Operator Theory in Engineering and Science", Holt, R. & W., 1971.

[21] MILNE-THOMSON L. M. - "The Calculus of Finite Differences", Macmillan, 1933.

[22] JORDAN C. - "Calculus of Finite Differences", Chelsea, 1950.

[23] NIELSEN K. L. - "Methods in Numerical Analysis", (2nd Ed.), Macmillan, 1964.

[24] BLUM E. K. - "Numerical Analysis and Computation: Theory and Practice", Addison-Wesley, 1972.

[25] HILDEBRAND F. B. - "Introduction to Numerical Analysis", (2nd Ed.), McGraw-Hill, 1974.

[26] ROBBINS H. and MONRO S. - "A Stochastic Approximation Method", Ann. of Math. Stat., Vol. 22, pp. 400-407, 1951.

[27] KIEFER J. and WOLFOWITZ J. - "Stochastic Estimation of the Maximum of a Regression Function", Ann. Math. Stat., Vol. 23, pp. 462-466, 1952.

[28] BLUM J. R. - "Multidimensional Stochastic Approximation Method", Ann. Math. Stat., Vol. 25, pp. 382-386, 1954.

[29] DVORETZKY A. - "On Stochastic Approximation", Proc. of the 3rd Berkeley Symp. on Math. Stat. & Prob., Vol. 1, (Ed. by J. Neyman), Univ. of California Press, pp. 39-55, 1956.

[30] ALBERT A. E. and GARDNER L. A., Jr. - "Stochastic Approximations and Nonlinear Regression", MIT Press, 1967.

[31] WASAN M. T. - "Stochastic Approximation", Cambridge Univ. Press, 1969.

[32] DERMAN C. - "Stochastic Approximation", Ann. Math. Stat., Vol. 30, pp. 879-886, 1956.

[33] WILDE D. J. - "Experimental Error", Chapter 6 of Optimum Seeking Methods, Prentice-Hall, 1964.

[34] SCHULTZ P. R. - "Some Elements of Stochastic Approximation Theory and Its Application to a Control Problem", in Modern Control Systems Theory, (Ed. by C. T. Leondes), McGraw-Hill, 1965.

[35] SAKRISON D. J. - "Stochastic Approximation: A Recursive Method for Solving Regression Problems", in Advances in Communication Systems, Vol. 2, (Ed. by A. V. Balakrishnan), Academic Press, 1966.

[36] FU K. S. - "Stochastic Approximation: A Brief Survey", Appendix F of Sequential Methods in Pattern Recognition and Machine Learning, Academic Press, 1968.

[37] FU K. S. - "Learning System Theory", in System Theory, (Ed. by L. A. Zadeh and E. Polak), McGraw-Hill, 1969.

[38] KASHYAP R. L., BLAYDON C. C. and FU K. S. - "Stochastic Approximation", in Adaptive, Learning, and Pattern Recognition Systems: Theory and Applications, (Ed. by J. M. Mendel and R. S. Fu), Academic Press, 1970.

[39] BLAYDON C. C., KASHYAP R. L. and FU K. S. - "Applications of the Stochastic Approximation Methods", in Adaptive, Learning, and Pattern Recognition Systems: Theory and Applications, (Ed. by J. M. Mendel and K. S. Fu), Academic Press, 1970.

[40] FABIAN V. - "Stochastic Approximation", in Optimizing Methods in Statistics, (Ed. by J. S. Rustagi), Academic Press, 1971.

[41] LJUNG L. - "Convergence of Recursive Stochastic Algorithms", Lund Inst. of Tech., Div. of Autom. Control, Report No. 7403, Feb. 1974.

[42] SARIDIS G. N. - "Stochastic Approximation Methods for Identification and Control - A Survey", IEEE Trans., Autom. Control, Vol. AC-19, No. 6, pp. 798-809, Dec. 1974.

[43] TSYPKIN Y. Z. - "Adaption, Treaning and Self Learning in Control Systems", Autom. & Rem. Control, Vol. 27, No. 1, pp. 16-51, Jan. 1966.

[44] KASHYAP R. L. and BLAYDON C. C. - "Recovery of Functions from Noisy Measurements Taken at Randomly Selected Points and it Application in Pattern Recognition", Proc. IEEE, Vol. 54, No. 8, pp. 1127-1129, Aug. 1966.

[45] SARIDIS G. N., NIKOLIC Z. J. and FU K. S. - "Stochastic Approximation Algorithms for System Identification, Estimation and Decomposition of Mixtures", IEEE Trans., Syst. Sci. & Cybernetics, Vol. SSC-5, No. 1, pp. 8-15, Jan. 1969.

[46] SAKRISON D. J. - "The Use of Stochastic Approximation to Solve the System Identification Problem", IEEE Trans., Autom. Control, Vol. AC-12, No. 5, pp. 563-567, Oct. 1967.

[47] SARIDIS G. N. and STEIN G. - "Stochastic Approximation Algorithms for Linear Discrete-Time System Identification", IEEE Trans., Autom. Control, Vol. AC-13, No. 5, pp. 515-523, Oct. 1968.

[48] SARIDIS G. N. and STEIN G. - "A New Algorithm for Linear System Identification", IEEE Trans., Autom. Control, Vol. AC-13, No. 5, pp. 592-594, Oct. 1968

[49] KUBRUSLY C. S. and GRAVIER J. P. - "Stochastic Approximation Algorithms and Applications", IEEE Dec. & Control Conf., San Diego, pp. 763-766, Dec. 1973.

[50] GRAUPE D. and PERL J. - "Stochastic Approximation Algorithm for Identifying ARMA Processes", Int. J. of System Sci., Vol. 5, No. 6, pp. 789-809, Dec. 1974.

[51] SAGE A. and MELSA J. L. - "System Identification", Academic Press, 1971.

[52] GRAUPE D. - "Identification of Systems", Van Nostrand, 1972.

[53] MENDEL J. M. - "Discrete Techniques of Parameter Estimation", Dekker, 1973.

[54] EYKHOFF P. - "System Identification: Parameter and State Estimation", Wiley, 1974.

[55] ZHIVOGLYADOV V. P. and KAIPOV V. K. - "Application of the Method of Stochastic Approximations in the Problem of Identification", Autom. & Rem. Control, Vol. 27, No. 10, pp. 1702-1706, Oct. 1966.

[56] ZHIVOGLYADOV V. P. and KAIPOV V. K. - "Identification of Distributed Plants in the Presence of Noises", IFAC Symp. on Ident. of Autom. Control Syst., Prague, p-3.5, Jun. 1967.

[57] ZHIVOGLYADOV V. P. and KAIPOV V. K. - "Accuracy of Distributed Systems Identification Algorithms", IFAC Symp. on Ident. & Proc. Param. Est., Prague, Jun. 1970.

[58] ZHIVOGLYADOV V. P., KAIPOV V. K. and TSIKUNOVA J. M. - "Stochastic Algorithms of Identification and Adaptive Control of Distributed Parameter Systems", IFAC Symp. on the Control of D.P.S., Banff, p-13.1, Jun. 1971.

[59] CARPENTER W. T., WOZNY M. J. and GOODSON R. E. - "Distributed Parameter Identification Using the Method of Characteristics", Trans. ASME, J. of Dyn. Syst., Meas. & Control, Vol. 93-G, No. 2, pp. 73-78, Jun. 1971.

[60] KUBRUSLY C. S. and CURTAIN R. F. - "Identification of Noisy Distributed Parameter Systems Using Stochastic Approximation", Control Theory Centre, Univ. of Warwick, CTC Report No. 37, May 1975, (to appear in Int. J. of Control).

[61] WOLFOWITZ J. - "On Stochastic Approximation Methods", Ann. Math. Stat. Vol. 27, pp. 1151-1156, 1956.

[62] DERMAN C. and SACKS J. - "On Dvoretzky's Stochastic Approximation Theorem", Ann. Math. Stat., Vol. 30, pp. 601-606, 1959.

[63] SCHMETTERER L. - "Sur L'Itération Stochastique", Le Calcul des Probabilites et ses Applications, Vol. 87, pp. 55-63, 1958.

[64] VENTER J. H. - "On Dvoretzky Stochastic Approximation Theorem", Ann. Math. Stat., Vol. 37, pp. 1534-1544, 1966.

[65] DOOB J. L. - "Stochastic Processes", Wiley, 1953.

[66] LOÈVE M. - "Probability Theory", (3rd Ed.), Van Nostrand, 1963.

[67] LUCKAS E. - "Stochastic Convergence", Heath-Raytheon, 1968.

[68] CHUNG K. L. - "A Course in Probability Theory", (2nd Ed.), Academic Press, 1974.

[69] STOUT W. F. - "Almost Sure Convergence", Academic Press, 1974.

[70] LUENBERGER D. G. - "Optimization by Vector Space Methods", Wiley, 1969.

[71] HIRSCH M. W. and SMALE S. - "Differential Equations, Dynamical Systems and Linear Algebra", Academic Press, 1974.

[72] BACHMAN G. and NARICI L. - "Functional Analysis", Academic Press, 1966.

[73] KNOPP K. - "Theory and Application of Infinite Series", (2nd Ed.), Blackie, 1951.

[74] FERRAR W. L. - "Convergence", Claredon Press-Oxford, 1938.

[75] WETHERILL G. B. - "Sequential Estimation of Points on Regression Functions", Chapter 9 of Sequential Methods in Statistics, (2nd Ed.), Chapman & Hall, 1975.

CHAPTER 4

[1] WONG E. - "Stochastic Processes in Information and Dynamical Systems", McGraw-Hill, 1971.

[2] CURTAIN R. F. and FALB P. L. - "Stochastic Differential Equations in Hilbert Space", J. of Differential Eq., Vol. 10, No. 3, pp. 412-430, Nov. 1971.

[3] CURTAIN R. F. - "Stochastic Parabolic Equations of Higher Order in t", J. of Math. Anal. & Appl., Vol. 46, No. 1, pp. 93-103, Apr. 1974.

[4] KUBRUSLY C. S. and CURTAIN R. F. - "Identification of Noisy Distributed Parameter Systems Using Stochastic Approximation", Control Theory Centre, Univ. of Warwick, CTC Report No. 37, May 1975 (to appear in Int. J. of Control).

[5] MEDITCH J. S. - "Stochastic Optimal Linear Estimation and Control", McGraw-Hill, 1969.

[6] SAGE A. P. and MELSA J. L. - "Estimation Theory with Application to Communication and Control", McGraw-Hill, 1971.

[7] MELSA J. L. and SAGE A. P. - "An Introduction to Probability and Stochastic Processes", Prentice-Hall, 1973.

[8] GANTMACHER F. R. - "The Theory of Matrices", Vol. I, Chelsea, 1960.

[9] JURY E. I. - "Inners and Stability of Dynamic Systems", Wiley, 1974.

[10] HOUSEHOLDER A. S. - "The Theory of Matrices in Numerical Analysis", Blaisdell, 1964.

[11] LUENBERGER D. G. - "Optimization by Vector Space Methods", Wiley, 1969.

CHAPTER 5

[1] KUBRUSLY C. S. and CURTAIN R. F. - "Identification of Noisy Distributed Parameter Systems Using Stochastic Approximation", Control Theory Centre, Univ. of Warwick, CTC Report No. 37, May 1975, (to appear in Int. J. of Control).

[2] COLLINS P. L. and KHATRI H. C. - "Identification of Distributed Parameter Systems Using Finite-Differences", Trans. ASME, J. of Basic Eng., Vol. 91-D, No. 2, pp. 239-245, Jun. 1969.

[3] ZHIVOGLYADOV V. P. and KAIPOV V. K. - "Application of the Method of Stochastic Approximations in the Problem of Identification", Autom. & Rem. Control, Vol. 27, No. 10, pp. 1702-1706, Oct. 1966.

[4] ZHIVOGLYADOV V. P. and KAIPOV V. K. - "Identification of Distributed Plants in the Presence of Noises", IFAC Symp. on Ident. of Autom. Control Syst., Prague, p-3.5, Jun. 1967.

[5] ZHIVOGLYADOV V. P. and KAIPOV V. K. - "Accuracy of Distributed Systems Identification Algorithms", IFAC Symp. on Ident. & Proc. Param. Est., Prague, Jun. 1970.

[6] ZHIVOGLYADOV V. P., KAIPOV V. K. and TSIKUNOVA J. M. - "Stochastic Algorithms of Identification and Adaptive Control of Distributed Parameter Systems", IFAC Symp. on the Control of D.P.S., Banff, p-13.1, Jun. 1971.

[7] CARPENTER W. T., WOZNY M. J. and GOODSON R. E. - "Distributed Parameter Identification Using the Method of Characteristics", Trans. ASME, J. of Dyn. Syst., Meas. & Control, Vol. 93-G, No. 2, pp. 73-78, Jun. 1971.

[8] TZAFESTAS S. G. - "Identification of Stochastic Distributed Parameter Systems", Int. J. of Control, Vol. 11, No. 4, pp. 619-624, 1970.

[9] TZAFESTAS S. G. - "Identification of Hybrid Distributed Parameter Systems", Int. J. of Control, No. 1, Vol. 13, pp. 145-154, Jan. 1971.

[10] BALAKRISHNAN A. V. - "Identification-Inverse Problems for Partial Differential Equations: A Stochastic Formulation", 6th IFIP Conf. on Opt. Tech., Lecture Notes in Computer Science, Vol. 27, Springer-Verlag, 1975.

[11] KUBRUSLY C. S. and GRAVIER J. P. - "Stochastic Approximation Algorithms and Applications", IEEE Dec. & Control Conf., San Diego, pp. 763-766, Dec. 1973.

[8] TZAFESTAS S. G. - "Identification of Stochastic Distributed Parameter Systems", Int. J. of Control, Vol. 11, No. 4, pp. 619-624, 1970.

[9] TZAFESTAS S. G. - "Identification of Hybrid Distributed Parameter Systems", Int. J. of Control, No. 1, Vol. 13, pp. 145-154, Jan. 1971.

[10] BALAKRISHNAN A. V. - "Identification-Inverse Problems for Partial Differential Equations: A Stochastic Formulation", 6th IFIP Conf. on Opt. Tech., Lecture Notes in Computer Science, Vol. 27, Springer-Verlag, 1975.

[11] KUBRUSLY C. S. and GRAVIER J. P. - "Stochastic Approximation Algorithms and Applications", IEEE Dec. & Control Conf., San Diego, pp. 763-766, Dec. 1973.