Identification of Physical Models

By
Henrik Melgaard

LYNGBY 1994
Ph.D. THESIS
NO. 4

IMM

ISSN 0909–3192

This documentation was prepared with LaTeX.
Printed by IMM-DTU
Bookbinder Hans Meyer

Preface

This thesis has been prepared at the IMSOR group of the Institute of Mathematical Modelling (IMM), the Technical University of Denmark, in partial fulfillment of the requirements for the degree of Ph.D. in engineering.

The thesis discuss the different aspects involved in the identification of a dynamical physical models. The work is focused on the use and incorporation of physical knowledge and other a priori information in all the phases involved in identification. Selected applications from modelling of building thermodynamics and of combustion engines are given.

The main contributions to this field is thought to be on experiment design for dynamical systems and on the implementation of the methods for estimation. A tool is developed and implemented, which is able to estimate the parameters of stiff physical models. Also the applications on the real systems represent new work in their respective fields.

Lyngby, June 1994

Henrik Melgaard

Acknowledgements

I wish to express my gratitude to a number of people, who have contributed to this research in one way or another.

First off all I would like to thank my supervisors, Assoc. Prof., Ph.D. Henrik Madsen, IMM, DTU, Techn. Dr. Jan Holst, Department of Mathematical Statistics, Lund Institute of Technology, Lund, Sweden and Assoc. Prof., Ph.D. Elbert L. Hendricks, Institute of Automatic Control Systems, DTU, for their help and guidance and for our inspiring discussions during this work.

I also want to thank my colleagues at IMSOR for a pleasant and constructive scientific and social environment. Furthermore I am grateful to the past and present members of the timeseries group at IMM for their collaboration.

The staff at the Laboratory for Energetics at DTU and acad. ing. Jan Lillelund and Elbert Hendricks are thanked for their support during data collection from a car engine at the Laboratory for Energetics.

I wish to thank the participants of the Commission of European Communities research project, PASSYS, for their cooperation on identification of thermal dynamics of buildings. Also the Danish participants from the Thermal Insulation Laboratory at DTU are thanked for their excellent collaboration.

Last, but not least I am grateful to my fiancée, Signe, and our children Emil and Karoline for their large patience during my preparation of this thesis.

Summary

The problem of identification of physical models is considered within the frame of stochastic differential equations. Methods for estimation of parameters of these continuous time models based on discrete time measurements are discussed. The important algorithms of a computer program for ML or MAP estimation of the parameters of nonlinear stochastic differential equations are described and the implemented tool is validated with respect to bias and uncertainty of the estimated parameters.

The different phases involved in identification of this type of models are considered in the thesis. This includes design of experiments, which is for instance the design of an input signal that are optimal according to a criterion based on the information provided by the experiment. Also model validation is discussed. An important verification of a physical model is to compare the physical characteristics of the model with the available prior knowledge.

The methods for identification of physical models have been applied in two different case studies. One case is the identification of thermal dynamics of building components. The work is related to a CEC research project called

PASSYS (Passive Solar Components and Systems Testing), on testing of building components related to passive solar energy conservation, tested under outdoor climate conditions.

The second case study is related to the performance of a spark ignition car engine. A phenomenological model of the fuel flow is identified under various operating conditions of the engine. This engine submodel is important for controlling the air/fuel ratio, e.g. in a feed-forward controller.

Resumé

Identifikation af fysiske modeller betragtes inden for rammerne a stokastiske differentialligninger. Metoder til estimation af parameterne i disse kontinuert tids modeller, baseret på diskrettids observationer, diskuteres. De vigtigste algoritmer i et computer program til ML eller MAP estimation af parameterne i ulinære stokastiske differentialligninger beskrives og det implementerede program-værktøj valideres med hensyn til bias og usikkerhed af de estimerede parametre.

De forskellige faser involveret i identifikation af denne type modeller behandles i afhandlingen. Dette inkluderer forsøgsdesign, hvilket for eksempel er design af et input signal, som er optimal med hensyn til et kriterium baseret på informationen fra forsøget. Model validering diskuteres ligeledes. En vigtig verificering af en fysisk model er at sammenligne fysiske karakteristika af modellen med den tilgængelige a priori viden.

Metoderne til identifikation af fysiske modeller er anvendt i to forskellige cases. Det ene case handler om identifikation af varmedynamikken af bygningskomponenter. Arbejdet er relateret til et EU forskningsprojekt, PASSYS (Passive Solar Components and Systems Testing), som omhandler

testning af bygningskomponenter til udnyttelse af passiv solvarme, testet under udendørs forhold.

Det andet case er relateret til ydeevnen af en benzinmotor. Der er identificeret en fænomenologisk model af benzinflowet i motoren under forskellige belastningsforhold af motoren. Denne undermodel af motoren er vigtig til kontrol af brændstof-luft forholdet, fx ved feed-forward kontrol af blandingsforholdet.

Contents

Chapter 1

Introduction

1.1 Background

The construction of mathematical models for dynamical systems from measured data, is termed *system identification*. The literature in this field is extensive, and some good introductions on the topic in general, can be found in (Goodwin & Payne, 1977; Ljung, 1987; Söderström & Stoica, 1989).

The identification of a certain model involves a number of steps in an iterative process. The first step is to decide the purpose of the model, e.g. should the model be used for controling the system, fault detection in the system or for determination of characteristic physical parameters etc. In any case, one of the important things is to make clear on which time scale the interesting dynamics of the system are. From these considerations some initial model is set up. Next step is to design an experiment for determination of the interesting parameters of the model. This experimental design is based on the intended use of the model and our prior knowledge about the system. From the observations provided by the experiment the parameters of the model structure are estimated. Then follows some iterative process of validating, and possibly modifying the model. If necessary another experiment is designed based on the new information gained and so on.

Based on the goal for modelling and the a priori knowledge it is usually possible to define the framework for the model, i.e. define the set of models, from which the solution of the identification problem is to be sought. Usually the definition of model structures are classified in two extremes according to their utilization of a priori knowledge about the process. The one extreme assumes that the model structure is known and deterministic. The parameters of the model are adjusted by minimizing the output error between simulated and measured output. The advantage by this modeling approach is that prior information is exploited. Often this means that the model structure is defined by the physical relations of the process (in general nonlinear), e.g. energy conservation, Newton's laws etc.

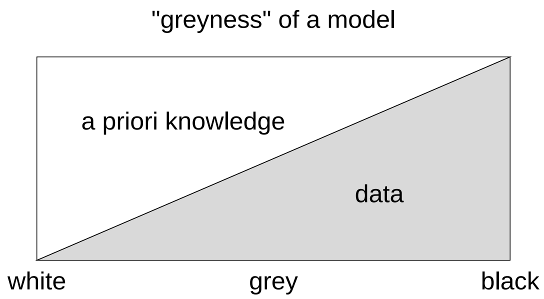
The opposite extreme is known as black box modeling. In this approach the model can be viewed as a box, which aims at describe the relations between the measured input/output data. A priori knowledge about the model structure is discarded and therefore the model and the parameters of it usually have little physical significance. Linear model structures have been widely used for black box modeling because they are mathematically attractive. This class of models is rich enough to cover a large number of applications, since the model is only viewed as a tool for approximating the data, thus selecting a model of sufficiently high order will often fit the data. The parameters of the model are estimated by statistical methods, which are also used for selecting and validating the models.

A third approach is a combination of the two extremes, and is called *grey box* modeling (Bohlin, 1984; Graebe, 1989). This approach exploits the a priori physical knowledge about the process, but the model structure and parameters are not assumed to be completely known. Hence the model is formulated in a stochastic framework as opposed to the white box approach. The parameters of the model are estimated as for black box models, by using statistically identification methods. These statistical tools are also used for validating and modifying the model. A typical grey box situation, is when the the model structure is determined by the physical relations of the process, hence the term *physical models*.

The main advantage of the grey box approach compared to the white box, is that the estimation is kept in a stochastic framework, hence statistical methods are available for model validation and structure modification etc. The advantage compared to the black box approach is the use of prior physical information about the process for structure determination, which in general is nonlinear. Because the grey box model is determined from the physical relations of the system, it is expected that the model is valid beyond the range covered by the measured data. Therefore it is also expected that a grey box model is able to make better long term predictions, than a black box model. The parameters of the grey box model have physical significance which is seldom the case for black box models.

An important difficulty about grey box models, is the estimation of their parameters. In general it is much more difficult to calculate the prediction errors, which is needed for the estimation, from a nonlinear stochastic model, than either from a nonlinear simulation model or from a linear stochastic model. That is one of the reasons for the rather limited use of grey box models in practical applications: the lack of available software tools for grey box identification. Other authors who have contributed in this field are (Bohlin, 1984, 1994; Graebe, 1990a, 1990b; Tulleken, 1993).

The grey box modelling approach, used here, is depicted in Fig. 1.1. It consist of different phases, which is also reflected in the layout of this work. The “greyness” of the resulting model is a matter of the weighting between the use of a priori knowledge and the information in data when building the model. If main weight is put on a priori knowledge the approach is merely white. On the other hand, if the main weight is on the data samples from the system, then it is called a black box approach.



The work presented in this thesis represents a grey box approach, where the available physical knowledge about the system as well as available data is used and incorporated in the iterative process of system identification. But since the a priori knowledge is only partial, measurements from the system are used to estimate the unknown / partially unknown parameters of the model. The incorporation of prior knowledge in the modelling, will influence all the steps of the classical identification process, see Fig. 1.1. On one hand it complicates the identification process, but hopefully the result

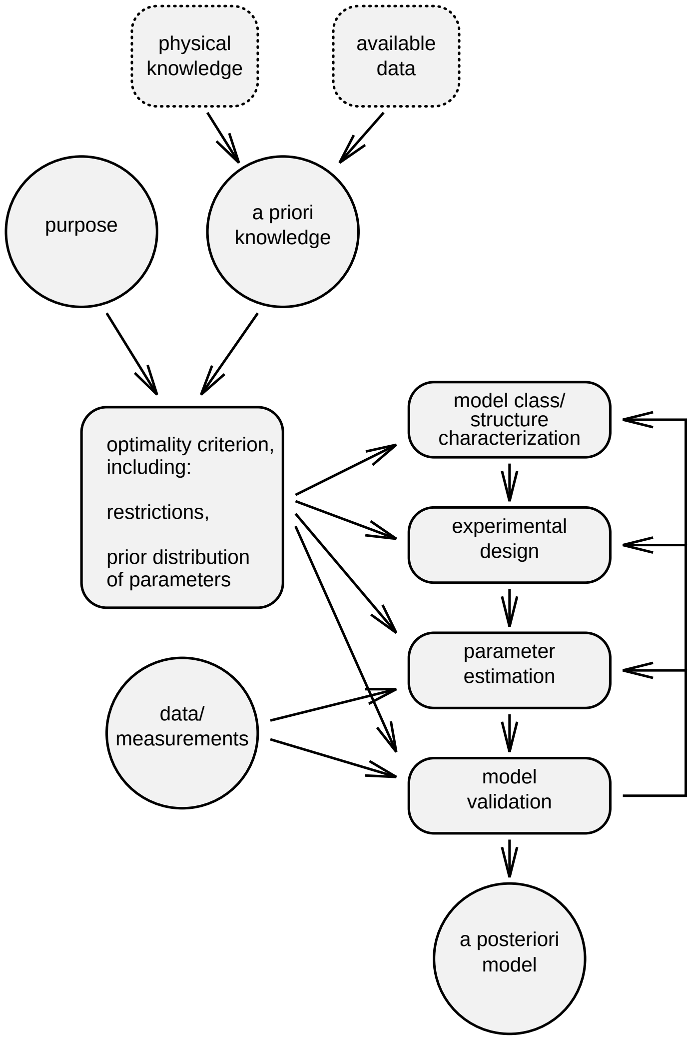


Figure 1.1. *The grey box modelling approach.*

is a better model, because it can be validated not only against measured data, but also against the prior knowledge about the system.

1.2 Outline of the thesis

In the present chapter the background and motivation for the work is given, and the organization of the thesis is outlined.

In Chapter 2 it is argued for that stochastic differential equations makes the obvious frame for modelling a wide range of physical systems. The mathematical and computational treatment of stochastic models is more complicated than of deterministic models. Simulation of the sample paths of stochastic differential equations is discussed in this chapter.

Chapter 3 is concerned with prediction error methods for parametric estimation of dynamical systems. Different asymptotic properties for the estimation procedures are discussed. This includes convergence and distribution of the parameters under certain conditions on the experimental conditions, the model, the criterion function etc. There has mainly been focused on the maximum likelihood estimator (ML) and on the maximum a posteriori estimator (MAP), because of their attractive property as asymptotically efficient estimator.

In Chapter 4 the general nonlinear filtering problem for a continuous time system, with discrete measurements is revealed. The conditional expected value and variance of the output is needed for the likelihood function for the parameter estimation. Different approximate solutions are discussed which are a compromise between the computational feasibility and the ability to handle nonlinear stochastic models.

The validation of physical models is discussed in Chapter 5. A number of different methods are outlined: tests for model structure, residual analysis, simulations and cross validation. An important verification of a physical model is to compare the physical characteristics of the model with the available prior knowledge.

Chapter 6 is concerned with the design of experiments. It is important for a good identification result that the experimental conditions are in accordance with the purpose of the model. The choice of input signal, sampling time and presampling filter are considered. Different measures of information all based on Fisher's information matrix are considered for the determination of optimal experiments. Also a bayesian approach of experimental design is discussed.

The methods for identification of physical models is used in two case studies. In Chapter 7 the methods have been applied for identification of the energy dynamics of building components. The work is related to a CEC (Commission of the European Communities) research project called PASSYS (Passive Solar Components and Systems Testing) for testing of building components related to passive solar energy conservation, tested under outdoor climate conditions.

Chapter 8 is a case study related to the performance of a spark ignition engine. A phenomenological model of the fuel flow is identified. This engine submodel is important for controlling the air/fuel ratio under transient conditions, and is therefore important for pollutant emissions and fuel economy. The work is performed on a 1.1L Ford engine with a central fuel injection.

In Appendix the implementation of the methods on a computer is discussed. The estimation of parameters in stochastic differential equations from discrete observations is not simple, and it is important to consider the numerical details of the calculations involved.

Chapter 2

Physical Models

Consider a real process or system. For some purpose we like to have a mathematical model of the system. The model should to a certain extend represent the system. If we have a good model then we expect the model to behave like the system when exposed to different manipulations. The different approaches for model building was briefly discussed in the previous chapter.

A typical grey box approach is based on physical modelling. This means that the physical knowledge of the system is used in the model. But unlike a white box model, the grey box model typically contains stochastic parts. There are a number of arguments for including such stochastic parts in the model. For one thing it represents the noise in the system, and noise is an intrinsic part of any real world system. This could for instance be the measurement noise when data are sampled from the system, using some measuring device. The stochastic part may also represent disturbances or inputs to the system which are not measured or known. It may further represent unmodelled dynamics of the system, e.g. the movements of the pistons in a meanvalue engine model, or turbulence in a flow. Often it is not directly the model of the noise we are interested in, the parameters of the noise model are considered as nuisance, but the noise model is still necessary in order to obtain a good description of the system. Usually though, some of the noise can be eliminated by filtering the data before using it for model estimation.

The inclusion of a stochastic part in the physical model complicates the mathematical treatment of the model. Thus, the next sections will define the stochastic model and describe tools to handle it.

2.1 Preliminaries

In this section there will be a very brief introduction to probability theory. The principal concept is the probability space $(\Omega, \mathcal{A}, P)$ consisting of a sample space Ω of possible outcomes, a σ-algebra $\mathcal{A}$ of subsets of Ω called events and a probability measure P, which assigns a probability $P(A)$ to each event A in $\mathcal{A}$. A family

$$\mathbf{x}(t, \omega), \qquad \omega \in \Omega, \qquad t \in J \tag{2.1}$$

of random variables is called a *stochastic process* with parameter set J and state space $\mathbb{R}^n$. For each fixed $t \in J$, $\mathbf{x}(t, \cdot)$ denotes a random variable on the probability space. For each fixed $\omega \in \Omega$, $\mathbf{x}(\cdot, \omega)$ corresponds to a $\mathbb{R}^n$-valued function defined on J. This is called a *sample path*, *trajectory*, or *realization* of the process.

A stochastic dynamical system satisfies the Markov property if the future state of the system at any time $t > s$ is independent of the past behavior of the system at times $t < s$, given the present state at time s. The stochastic process generated by this system is called a *Markov process*. The Markov property in terms of density functions, can be written

$$p(\mathbf{x}_t|\mathcal{A}_s^x) = p(\mathbf{x}_t|\mathbf{x}_s) \tag{2.2}$$

for $0 \le s \le t$, where $\mathcal{A}_s^x$ is defined as the σ-algebra generated by $\mathbf{x}$ in the time interval $[0, s]$. The subscript is the time index, and the probability space variable ω is suppressed in the notation. The increasing family $\{\mathcal{A}_t, t \ge 0\}$ of sub-σ-algebras of $\mathcal{A}$, i.e. $\mathcal{A}_s \subset \mathcal{A}_t \subseteq \mathcal{A}$ for $0 \le s < t$ is called a *filtration*. It may be interpreted as the σ-algebra containing all information about the involved processes in $[0, t]$. The particular σ-algebra generated by the stochastic process $\{\mathbf{x}_\tau, \tau \in [0, t]\}$ is denoted $\mathcal{A}_t^x$, and correspondingly the sequence $\{\mathcal{A}_t^x, t \ge 0\}$ is the filtration generated by $\mathbf{x}$.

The conditional densities $p(s, \mathbf{x}, t, \mathbf{y}) = p(\mathbf{x}_t = \mathbf{y}|\mathbf{x}_s = \mathbf{x})$ are called the transition probability densities of the Markov process. A Markov process $\mathbf{x}_t$ with transition densities $p(t, \mathbf{x}, s, \mathbf{y})$ is called a *diffusion process* if the following three limits exists for every $s \ge 0$, $\mathbf{x} \in \mathbb{R}^n$, and $\epsilon > 0$:

$$\lim_{t \downarrow s} \frac{1}{t-s} \int_{|y-x|>\epsilon} p(s, \mathbf{x}, t, \mathbf{y})\, d\mathbf{y} = 0 \tag{2.3}$$

$$\lim_{t \downarrow s} \frac{1}{t-s} \int_{|y-x|\le\epsilon} (\mathbf{y} - \mathbf{x}) p(s, \mathbf{x}, t, \mathbf{y})\, d\mathbf{y} = \mathbf{a}(s, \mathbf{x}) \tag{2.4}$$

$$\lim_{t \downarrow s} \frac{1}{t-s} \int_{|y-x|\le\epsilon} (\mathbf{y} - \mathbf{x})(\mathbf{y} - \mathbf{x})^T p(s, \mathbf{x}, t, \mathbf{y})\, d\mathbf{y} = \mathbf{b}(s, \mathbf{x}) \tag{2.5}$$

where $\mathbf{a}$ and $\mathbf{b}$ are well defined functions. The quantity $\mathbf{a}(s, \mathbf{x})$ is called the drift of the diffusion process and $\mathbf{b}(s, \mathbf{x})$ its diffusion matrix. $\mathbf{b}$ is symmetric and positive semidefinite. Condition (2.3) prevents a diffusion process from having instantaneous jumps. Properties (2.4) and (2.5) can be written,

$$E(\mathbf{x}_t - \mathbf{x}_s | \mathbf{x}_s = \mathbf{x}) = \mathbf{a}(s, \mathbf{x})(t-s) + o(t-s) \tag{2.6}$$
$$E((\mathbf{x}_t - \mathbf{x}_s)(\mathbf{x}_t - \mathbf{x}_s)^T | \mathbf{x}_s = \mathbf{x}) = \mathbf{b}(s, \mathbf{x})(t-s) + o(t-s) \tag{2.7}$$

so the drift gives the rate of change in the conditional mean of the process. The diffusion matrix represents the rate of change of the conditional covariance of the increment.

Solutions to stochastic differential equations, to be discussed in later sections, are Markov diffusion processes. The standard Wiener process, see the next section, is a diffusion process with $\mathbf{a} = 0$ and $\mathbf{b} = \mathbf{I}$, and is the solution of the simplest stochastic differential equation.

A final concept to bring up in this section is that of *martingale*. A stochastic process $\{\mathbf{x}_t, t \geq 0\}$ adapted to the filtration $\mathcal{A}_t$ is called a martingale if $E(|\mathbf{x}_t|) < \infty$, and for all $0 \leq s < t$,

$$E(\mathbf{x}_t | \mathcal{A}_s) = \mathbf{x}_s \quad \text{w.p.1} \tag{2.8}$$

The simplicity of predicting a martingale according to (2.8) explain the importance of this concept for applications. Such a process is sometimes referred to as “fair game” processes; if $\mathbf{x}_s$ represents a gampler’s fortune at time s, the game is fair if his expected fortune at a future time $t > s$ given the game history up to some previous time s is precisely the fortune at time s. If the equality sign in (2.8) is replaced by a $\leq$ then the process is called a *supermartingale*, and similarly a *submartingale* if it is replaced by a $\geq$. An example of a martingale is the Wiener process. A more thorough discussion of the different topics introduced in this section may be found in e.g. (Doob, 1953; Gihman & Skorohod, 1979; Karatzas & Shreve, 1988).

2.1.1 Brownian Motion

Brownian movement is the name originally given to the irregular movement of pollen, suspended in water, observed by the botanist Robert Brown in 1828. This random movement, usually attributed to the buffeting of the pollen by water molecules, results in a diffusion of the pollen in the water. A standard Wiener process is a fundamental stochastic process providing a mathematical description of the physical process of Brownian motion. The range of application of this process goes far beyond a study of microscopic particles in suspension and includes modeling of noise and perturbations in thermal, electrical, biological and economical systems etc.

The mathematical properties defining a Wiener process, $\{\beta_t, t \geq 0\}$, are

(i) $\beta_0 = 0$ w.p.1

(ii) The increments $\beta_1 - \beta_0$, $\beta_2 - \beta_1$, $\cdots$, $\beta_n - \beta_{n-1}$, of the process, for any partitioning of the time interval $0 \leq t_0 < t_1 < \cdots < t_n < \infty$ are mutually independent.

(iii) The increment $\beta_t - \beta_s$ for any $0 \leq s < t$ is Gaussian with mean and covariance respectively

$$E(\beta_t - \beta_s) = 0\,; \qquad V(\beta_t - \beta_s) = \mathbf{I}|t - s| \tag{2.9}$$

here the *standard* Wiener process is defined by using the identity matrix in (2.9) instead of a general positive definite matrix.

There are a number of other important properties which characterizes a Wiener process, $\{\beta_t, t \geq 0\}$ adapted to the filtration $\mathcal{A}_t$. The Wiener process is both Markov and a martingale with respect to $\mathcal{A}_t$. The sample paths of the process is continuous with probability one, but they are nowhere differentiable w.p.1. Since the sample paths are almost surely continuous functions of time and the variance of the process grows unbounded as time increases, while the mean remains zero, according to (2.9), there must be sample paths attaining larger and larger (absolute) values as time increases. By using the strong law of large numbers one finds

$$\lim_{t \to \infty} \frac{\beta_t}{t} = 0 \tag{2.10}$$

More precise statements about the asymptotic behavior is given by the law of the iterated logarithm, which says that

$$\limsup_{t \to \infty} \frac{|\beta_t|}{\sqrt{2t \log \log t}} = 1 \; ; \qquad \liminf_{t \to \infty} \frac{|\beta_t|}{\sqrt{2t \log \log t}} = -1 \tag{2.11}$$

Equations (2.10) and (2.11) are all valid with probability one.

Even though the Wiener process is not differentiable, one may consider its time derivative $\dot{\beta}_t$, called *Gaussian white noise.* Mathematically this only make sense as a generalized function. The term *white* comes from the fact that the process has a uniform spectral density function, $f(\nu)$, for all frequencies $\nu \in \mathbb{R}$, which is a characteristic of white light. The spectral density is given by the Fourier transform of the covariance function, $\gamma(t)$. In the scalar case this is

$$f(\nu) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\nu t} \gamma(t)\, dt \,, \qquad \nu \in \mathbb{R} \tag{2.12}$$

By setting the spectral density to a constant $f(\nu) = c/(2\pi)$, for all frequencies, the covariance function satisfy formally

$$\gamma(t) = c\, \delta(t) \tag{2.13}$$

for all t, where $\delta(t)$ is the Dirac delta function. The nature of the covariance function (2.13) indicates that such a process cannot be realized physically, and that (continuous time) white noise is not a stochastic process in the usual sense. It can however be approximated to any desired degree of accuracy by conventional stochastic processes with broad banded spectra, such as the Ornstein-Uhlenbeck process, see (Gard, 1988).

2.1.2 Stochastic Integrals

In order to be able to handle stochastic differential equations it is necessary to use other integral concepts than the standard Riemann integral. Consider the following scalar stochastic differential equation

$$dx_t = a(t, x_t)\,dt + b(t, x_t)\,d\beta_t \tag{2.14}$$

where a is the drift coefficient and b is the diffusion coefficient, and β_t is a standard Wiener process, representing the source of noise in the system. A solution to (2.14) would have the form

$$x_t(\omega) = x_0(\omega) + \int_0^t a(s, x_s(\omega))\,ds + \int_0^t b(s, x_s(\omega))\,d\beta_s(\omega) \tag{2.15}$$

for each $\omega \in \Omega$. As it was shown in the previous section the sample paths of the Wiener process has unbounded variation so the Riemann integral of the last term will not converge. An appropriate stochastic integral to be defined is the *Itô stochastic integral*. This integral is defined as the mean-square limit of the left hand rectangular approximation

$$I(b) \triangleq \sum_{i=0}^{N-1} b(t_i, x(t_i, \omega))(\beta_{i+1}(\omega) - \beta_i(\omega)) \tag{2.16}$$

for all partitions $t_0 < t_1 < \cdots < t_N = t$ as the maximum step size $\delta = \max_i(t_{i+1} - t_i) \to 0$. An existence and uniqueness theorem, proved using successive approximations, holds for (2.16) when the drift and diffusion coefficients satisfy Lipschitz and bounded growth conditions, see (Gard, 1988; Kloeden & Platen, 1992). The limit of (2.16) is called the stochastic integral in the sense of Itô. The solution process $x_t(\omega)$ is a Markov diffusion process. An important difference between classical calculus and Itô calculus occurs in their transformation or chain rules. For $y_t = f(t, x_t)$ with x_t a solution of (2.14) and f a sufficiently smooth function we obtain

$$dy_t = (\frac{\partial f}{\partial t} + a\frac{\partial f}{\partial x} + \frac{1}{2}b^2\frac{\partial^2 f}{\partial x^2})\,dt + b\frac{\partial f}{\partial x}\,d\beta_t \tag{2.17}$$

where all the terms on the right side are evaluated at (t, x_t). Equation (2.17), which is known as the *Itô formula*, contains an additional term $\frac{1}{2}b^2\frac{\partial^2 f}{\partial x^2}$ which would not be present if the rules of classical calculus held, see (Gard, 1988).

Other stochastic integrals have also been proposed. The most important of these is the *Stratonovich integral*. It evaluates the integrand of the stochastic integral at the midpoint of each partition subinterval rather than at the left hand point as in (2.16). Hence the Stratonovich integral is defined as the mean-square limit of the approximation

$$S(b) \triangleq \sum_{i=0}^{N-1} b(\tau_i, x(\tau_i, \omega))(\beta_{i+1}(\omega) - \beta_i(\omega)) \tag{2.18}$$

for $\tau_i = (t_i + t_{i+1})/2$ as $\delta \to 0$. The Stratonovich integral obeys the transformation rules of classical calculus, i.e. it does not contain the additional term of the Itô formula, which is a major reason for its use. Stochastic processes defined by Stratonovich integrals do not satisfy martingale and Markov properties of their Itô integral counterparts.

Fortunately, there is a connection between the two integrals, with the solution of (2.14) in the Itô sense (2.16) being similar to the solution in the Stratonovich sense with the drift coefficient of (2.14) changed to

$$a'(t, \mathbf{x}_t) = a(t, \mathbf{x}_t) - \frac{1}{2} b(t, \mathbf{x}_t) \frac{\partial b(t, \mathbf{x}_t)}{\partial x} \tag{2.19}$$

and similarly for a transformation in the other direction. Thus once one of the calculi has been decided on, the advantages of the other can be exploited by means of this simple modification. It should be noted, that if the diffusion coefficient is only dependent on the time, then the solutions by the two integrals coincide.

2.2 Stochastic Differential Equations

Ordinary differential equations which have the general form

$$d\mathbf{x}_t = \mathbf{f}(t, \mathbf{x}_t)\, dt\,, \qquad t \geq 0 \tag{2.20}$$

provide deterministic descriptions of e.g. the laws of motion of physical systems. But usually we a operating under noise levels that are too large to be ignored, i.e. averaged out in the deterministic model. There are different approaches of how the noise of the system should be described by the model. The first, and simpler, class arises when an ordinary differential equation has random coefficients, a random initial value or is forced by a fairly regular stochastic process, for which the solution processes have differentiable sample paths. The equations are called *random differential equations* and are solved sample path by sample path as ordinary differential equations. The second class occurs when the forcing is an irregular stochastic process such as Gaussian white noise. The equations are then written symbolically as stochastic differentials, but are interpreted as integral equations with stochastic integrals, e.g. the Itô integral. They are called *stochastic differential equations*. Here we have chosen to consider models of the second class in the form of the *Itô equation*

$$d\mathbf{x}_t = \mathbf{f}(t, \mathbf{x}_t)\, dt + \mathbf{G}(t, \mathbf{x}_t) d\beta_t\,, \qquad t \geq 0 \tag{2.21}$$

where β_t is a n-dimensional standard Wiener process. As it was seen in the previous section the solution $\mathbf{x}_t$ of the Itô equation is a Markov diffusion process. Hence the transition probability densities $p(s, \mathbf{x}, t, \mathbf{y})$ of the solution of (2.21), described in section 2.1, satisfies *Kolmogorov's forward equation* or the *Fokker-Planck equation*, which is discussed in Chapter 4.

2.2.1 Physical and Mathematical Noise

There may be a problem of motivating that the process (2.21) is a good choice for a model for physical reality. Let us consider a physical noise η_t. Being a physical quantity we would expect both that η_t itself is absolutely continuous (which is satisfied by the Wiener process), and that it has bounded derivative. Otherwise the physical signal it represents could change its value discontinuously and infinitely fast. The physically reasonable properties of absolute continuity and bounded derivative imply Lipschitz continuity, i.e. there exists a constant K such that for all t, s we have $|\eta_t - \eta_s| \leq K|t - s|$. But the Brownian motion or Wiener process considered in (2.21) is almost surely non-differentiable and non-Lipschitzian, thus placing us in the dilemma that physical noise is Lipschitzian, whereas the mathematical abstraction we would like to employ is non-Lipschitzian. The Wiener process as the forcing process is tractable from a mathematical point of view because of its Markov and martingale properties. The problem is how to keep the Wiener process as the tool for the mathematical analysis, but accept that physical noise is Lipschitzian. Graebe (1990b) discusses different approaches to this problem, but here we just consider the use of shaping filters.

Assume that we have modelled our system using quantities of the system state $\mathbf{x}_t^s$, a physical noise vector η_t to form the differential equation

$$d\mathbf{x}_t^s = \mathbf{f}^s(t, \mathbf{x}_t^s, \eta_t)\, dt \tag{2.22}$$

Since η_t is Lipschitzian, (2.22) is simply an ordinary differential equation for every sample path of η_t. The idea of a shaping filter now is to recognize the physical noise properties of smoothness and bounded variation are also shared by the output of an Itô equation. One might therefore investigate the existence of an Itô equation that is driven by the Wiener process and that has η_t as its solution. Such an equation, is called a shaping filter since it shapes the spectrum of the Wiener process to the desired physical noise. Assume that we have found a shaping filter of the form

$$d\eta_t = \mathbf{f}^f(t, \eta_t)\, dt + \mathbf{G}^f(t, \eta_t) d\beta_t \tag{2.23}$$

where the superscript f denotes the filter equations, and β_t is a standard Wiener process. We now consider the augmented state, $\mathbf{x}_t$, of the system state, $\mathbf{x}_t^s$, and filter state, η_t, resulting in the definitions

$$\mathbf{x}_t \triangleq \begin{bmatrix} \eta_t \\ \cdot\cdot \\ \mathbf{x}_t^s \end{bmatrix}, \quad \mathbf{f} \triangleq \begin{bmatrix} \mathbf{f}^f(t, \eta_t) \\ \cdots\cdots\cdots \\ \mathbf{f}^s(t, \mathbf{x}_t^s, \eta_t) \end{bmatrix}, \quad \mathbf{G} \triangleq \begin{bmatrix} \mathbf{G}^f(t, \eta_t) \\ \cdots\cdots \\ 0 \end{bmatrix}. \tag{2.24}$$

Hence we have rewritten the original equation (2.22) into the form

$$d\mathbf{x}_t = \mathbf{f}(t, \mathbf{x}_t)\, dt + \mathbf{G}(t, \mathbf{x}_t) d\beta_t \tag{2.25}$$

which is again an Itô equation, though of higher order than originally. Using these arguments, we conclude that the Itô equation is able to represent a wide range of physical models. Note that in the model structure (2.22) the physical noise is not necessarily additive as is the Wiener process in the Itô equation.

2.3 Qualitative Analysis

Explicit solutions for stochastic differential equations are in general not possible to obtain. The qualitative theory of stochastic differential equations permits investigating the general behavior of solutions directly from the form of the differential equation. The qualitative theory contains topics as boundedness, stability, and uniqueness of solutions of stochastic differential equations. Modern differential equations theory, much of which is qualitative theory, had its beginnings at the end of the last century with the work of Poincaré on celestial mechanics and Lyapunov's study of the stability of motions.

In the previous section it was mentioned that Lipschitz conditions on $\mathbf{f}$ and $\mathbf{G}$ suffice to guarantee the existence and uniqueness of the solution of (2.21), see (Gard, 1988).

The counterpart of a deterministic equilibrium in a stochastic system is a stationary solution, $\bar{\mathbf{x}}_t$, which has a probability distribution that does not depend on time. Lyapunov functions provide one means of investigating the stability of such stationary solutions (Kushner, 1967). When linearizing (2.21) about a stationary solution, $\bar{\mathbf{x}}_t$, we obtain a linear stochastic differential equation, with $\boldsymbol{z}_t = \mathbf{x}_t - \bar{\mathbf{x}}_t$

$$d\boldsymbol{z}_t = \mathbf{f}'(t, \bar{\mathbf{x}}_t)\boldsymbol{z}_t \, dt + \mathbf{G}'(t, \bar{\mathbf{x}}_t)\boldsymbol{z}_t \, d\beta_t \tag{2.26}$$

where $\mathbf{f}'$ denotes $\partial\mathbf{f}/\partial\mathbf{x}$ etc. The exponential rate of convergence or divergence of a solution $\boldsymbol{z}_t$ of (2.26) from the null solution

$$\lambda(\boldsymbol{z}_0) = \lim_{t\to\infty} \sup \frac{1}{t} \log |\boldsymbol{z}_t| \tag{2.27}$$

is known as a *Lyapunov exponent* and play the same role as the real part of an eigenvalue in deterministic systems. The asymptotic stability of the null solution of the linear stochastic differential equation, and consequently the stationary solution of the original equation, is thus characterized by the negativity of the largest Lyapunov exponent λ_1. Lyapunov exponents provide an indication of the time scales in a dynamical system. Consider a system with Lyapunov exponents $\lambda_d \leq \cdots \leq \lambda_2 \leq \lambda_1$ and d is the dimension of the problem, if specifically

$$\lambda_d \ll \lambda_1 \tag{2.28}$$

there is a vastly differing of time scales, and the system is said to be stiff. The problem is that, unlike the eigenvalues of a deterministic system, the Lyapunov exponents of a stochastic system are very difficult to evaluate explicitly. Numerical approximations have been derived to calculate the top Lyapunov exponent λ_1. For instance Talay (1991) has proposed such a method, which uses simulations of approximate trajectories of a system to evaluate an approximation of its upper Lyapunov exponent.

2.4 Applications of Stochastic Differential Equations

In many cases a model of a physical system is specified by deterministic ordinary differential equations. There may then be a need to take into account random phenomena explicitly in order to obtain the desired accuracy in the modelling. The random aspects considered may be intrinsic, for example the mechanical noise from bearings in an engine, or external, such as random environmental characteristics affecting the drivability of a car on the road. Markov diffusion processes, which can be represented as solutions of stochastic differential equations, arise as tractable approximations to these stochastic model quantities.

In this section a selection of examples from the literature of applications of stochastic differential equations is given, to show the variety of disciplines where stochastic differential equation have been used. Here examples from modelling of population dynamics in biological systems given in (Gard, 1988, chapter 6) are shown. An example of an application in economics, following (Karatzas & Shreve, 1988, section 5.8), modelling investment/consumption theory is also given. A number of other examples from applications of stochastic differential equations may be found in (Kloeden & Platen, 1992, chapter 7) which presents examples from a wide range of fields, including biology, economics and different branches of physics. In chapter 7 and 8 in this dissertation two cases, using stochastic differential equations to model the thermal dynamics of buildings and the energy flows of a car engine respectively are shown, see also (Melgaard, Hendricks, & Madsen, 1990; Madsen, Melgaard, & Holst, 1990; Melgaard, Madsen, & Holst, 1992a).

In this section we do not enter into detailed discussions about the exact use of the models, but they are in general mainly used for qualitative investigations and for simulating and predicting the sample paths of the processes. If the parameters in the models are to be estimated on the basis of measurements from the system, the requirements for computational effort are high. The modelling aims in the cases considered in chapter 7 and 8 were parameter estimation in the models based on discrete time observations. Similar examples may be found in (Graebe, 1990a).

2.4.1 Population Dynamics

The simplest population dynamics models take the form of the differential equation $dx_t/dt = rx_t$, where r is a constant, representing the growth rate of the species modeled. The random environmental effects on the population can be modelled by replacing the growth rate by $\rho + \sigma\xi_t$ for a Gaussian white noise process, ξ_t. Thus resulting in a stochastic differential equation. Both the deterministic and stochastic model exhibit unbounded growth, which is untenable in an environment with finite resources. Under such circumstances a finite supportable carrying capacity K is appropriate, with the population decreasing whenever it exceeds this value. This results in the deterministic logistic or *Verhulst model*

$$dx_t/dt = rx_t(1 - x_t/K) \tag{2.29}$$

where r is the intrinsic growth rate of the population. By replacing the growth rate by $\rho+\sigma\xi_t$ as before, we obtain the following stochastic logistic model

$$dx_t = \rho x_t(1 - x_t/K)\,dt + \sigma x_t((1 - x_t/K)\,d\beta_t \tag{2.30}$$

Usually the single species population dynamics models are unrealistic since in nature most species coexist with others and are affected by their presence one way or another. The *Lotka-Volterra system* of ordinary differential equations

$$dx_t^i/dt = x_t^i(a^i + \sum_{j=1}^{n} b^{ij} x_t^j) , \qquad 1 \leq i \leq n \tag{2.31}$$

constitutes a simple nonlinear model of interacting multispecies population dynamics. In (2.31), the intrinsic growth rates a^i, and the interaction rates b^{ij} are assumed, in the simplest case, to be constants whose signs indicate whether the model represents prey-predator, competition, mutualism, or some mixture of these population dynamics types. Randomizing the growth rates a^i as $a^i + \sigma^i \xi_t^i$ leads to a system of stochastic differential equations

$$dx_t^i = x_t^i(a^i + \sum_{j=1}^{n} b^{ij} x_t^j)\, dt + \sigma^i x_t^i \, d\beta_t^i \tag{2.32}$$

which is a multidimensional Itô equation. There are other possibilities for parameterizing the stochastic models, refer to (Gard, 1988).

2.4.2 Investment Finance

Stochastic differential equations have been used in continuous time modelling of the trading of risky securities. Merton (1971) has formulated the problem of optimal consumption/investment in this framework. He considered an investor who chooses between two types of assets, one is safe and the other is risky. One of the assets, called the *bond*, has a price p_t^b

which evolves according to the differential equation

$$dp_t^b = r_t\, p_t^b\, dt\,, \qquad 0 \le t \le T \tag{2.33}$$

where $\{r_t; 0 \le t \le T\}$ is called the *interest rate* process. The other assets, called *stocks*, are risky, and their prices are modeled by the stochastic differential equation

$$dp_t^s = b_t\, p_t^s\, dt + \sigma_t p_t^s\, d\beta_t\,, \qquad 0 \le t \le T \tag{2.34}$$

where β_t is a standard Wiener process, b_t is called the *mean rate of return*. At each instant of time the investor must select the fraction f_t of his available capital or *wealth* that he will put into the risky investment, with the remaining fraction $1 - f_t$ going into the safe one. By combining (2.33) and (2.34) and assuming that his current *consumption rate* is $c_t \ge 0$, it follows that his wealth, x_t satisfies the following equation

$$dx_t = f_t(b_t\, x_t\, dt + \sigma_t x_t d\beta_t) + (1 - f_t)r_t\, x_t\, dt - c_t\, dt \tag{2.35}$$

which can be rewritten as the following Itô equation

$$dx_t = ((f_t b_t + (1 - f_t)r_t)x_t - c_t)dt + f_t\sigma_t x_t\, d\beta_t \tag{2.36}$$

When the investor has perfect information about his current wealth, feedback controls of the form $(f_t, c_t)^T = \mathbf{u}(t, x_t)$ provide a natural way for choosing his current investment mixture and consumption rate. The investor wishes to choose $\mathbf{u}$, so as to maximize the expected value of some utility function U at time T. This formulates an optimal stochastic control problem with profit functional

$$J(s, x, \mathbf{u}) = E(U(x_T^u) | x_s^u = x) \tag{2.37}$$

to be maximized, where $\{x_t^u; 0 \leq t \leq T\}$ is the solution of (2.36) including the feedback law. The problem is complicated by the presence of a non-negative consumption rate, which may result in bankruptcy at a random first exit time

$$\tau = \inf(t \geq s : x_t^u = 0 | x_s^u = x) \tag{2.38}$$

If $\tau < T$ we say that *bankruptcy* occurs at time τ. The example of an application in economics given here follows Karatzas & Shreve (1988, section 5.8), where further examples may be found.

2.5 Simulation of Stochastic Differential Equations

Explicit solutions of stochastic differential equations are only possible for simple linear equations. In general one has to resort to some numerical approximation of the solutions. Different numerical approaches have been proposed (Gard, 1988; Kloeden & Platen, 1989), such as Markov chain approximations where both the state and time variables are discretized. Here we focus on time discrete approximations as they usually are effective for a wider range of situations.

Simulating a nonlinear stochastic differential equation is usually easier and requires less approximations than computing conditional expectations for the same process. In principle it involves running sequences from a random number generator through the discretized equation. Hence, we obtain approximations to one or several sample paths of the solution. Simulating the sample paths is an important tool both for applications in qualitative analysis (Section 2.3) as for the visual validation of a model. If one is also interested in the statistical properties of the solution, then a whole family of sample paths of the solution is needed. This method is called Monte Carlo simulation and it is able to give the evolution of the whole distribution of the solution.

The solution of a (scalar) stochastic differential equation (2.14) is formulated as an integral equation

$$x_t(\omega) = x_0(\omega) + \int_0^t a(s, x_s(\omega))\, ds + \int_0^t b(s, x_s(\omega))\, d\beta_s(\omega) \qquad (2.39)$$

where the first integral is a standard Riemann integral for each $\omega \in \Omega$ and the second is an Itô stochastic integral defined in Section 2.1.2. Consider a time discretization of the time interval $[0, T]$ given as

$$0 = t_0 < t_1 < \cdots < t_n = T \qquad (2.40)$$

with step size $\Delta t_i = t_{i+1} - t_i$. The maximum step size is $\delta = \max \Delta t_i$. A discrete time approximation of a solution x_t of (2.21) is a sequence $\{y_k\}$, with y_i approximating x_t at time $t = t_i$. The simplest approximation of a stochastic differential equation (2.21) is the *Euler approximation*. It has the form

$$y_{i+1} = y_i + a(t_i, y_i)\, \Delta t_i + b(t_i, y_i)\, \Delta\beta_i \qquad (2.41)$$

for $i = 0, 1, \cdots, n$ and initial value $y_0 = x_0$. The noise increments are given by $\Delta\beta_i = \beta_{t_{i+1}} - \beta_{t_i}$. Equation (2.41) is derived by fixing the integrands of both integrals in (2.39) to the left end point of each discretization interval, which corresponds to the definition of the Itô integral. If values are required at intermediate instants either piecewise constant values from the preceding discretization point or some interpolation, e.g. linear, of the values at the two immediate enclosing discretization times could be used. The standard Wiener increments $\Delta\beta_i$ in (2.41) are $N(0, \Delta t_i)$ distributed random variables. They are easily generated by a pseudo random number generator, see Section A.6.

2.5.1 Stochastic Taylor Expansion

In the previous section only the simple Euler integration was considered, but also higher order approximations are possible for stochastic integrals. There are different criteria for grouping these methods according to their properties. If y_T^δ is a time discrete approximation at the terminal time $t = T$ with a maximum step length δ, it is said to *converge strongly with order* $\lambda > 0$ if there exists some positive constant C independent of δ

$$E|y_T^\delta - x_T| \leq C\delta^\lambda \tag{2.42}$$

for all $\delta \in]0, \delta_0[$ where $\delta_0 > 0$. This criterion measures the absolute error of the approximation at the terminal time. Another frequently used criterion for strong convergence is by the quadratic mean squared expression

$$E(|y_T^\delta - x_T|^2), \tag{2.43}$$

related to the largest increment δ. The expression (2.43) measures the *global* error over $t \in [0; T]$. Expressions based on one-step errors measures the *local* error. If only an approximation of the probability distribution is required the closeness of moments is of interest and not the sample paths themselves. The time discrete approximation *converges weakly with order* β if

$$|E(g(y_n^\delta)) - E(g(x_T))| \leq C\delta^\beta \tag{2.44}$$

for some function g which is continuously differentiable of sufficiently high order. This definition implies the convergence of all moments. It turns out that under appropriate smoothness conditions the Euler approximation (2.41) converges strongly with order $\lambda = 0.5$ and weakly with order $\beta = 1.0$, see Kloeden & Platen (1992).

Another way of classifying the different methods is to compare them with a truncated Taylor expansion. There are several possibilities for such a Taylor expansion for a stochastic integral. One is based on iterated application of the Itô formula (2.17), which is called the *Itô-Taylor expansion* following Kloeden & Platen (1992). Consider a smooth function of an Itô process $f(x_t)$ to be expanded about $f(x_0)$, for the scalar case

$$\begin{aligned} f(x_t) = f(x_0) &+ (af' + \tfrac{1}{2}b^2 f'')|_{x_0} \int_0^t ds + (bf')|_{x_0} \int_0^t d\beta_s \\ &+ (b(bf')')|_{x_0} \int_0^t \int_0^s d\beta_u d\beta_s + R \end{aligned} \tag{2.45}$$

where quote denotes partial derivative with respect to x. The remainder terms R involve higher order multiple stochastic integrals with variable integrands. The expansion (2.45) has only been developed to one extra term compared to first order. Applying such expansions over each time discretization interval with $f(x_t) \equiv x_t$ and truncating (by neglecting R in 2.45), we obtain the *Milshtein approximation*, (Milshtein, 1974)

$$\begin{aligned} y_{i+1} = y_i &+ a(t_i, y_i)\,\Delta t_i + b(t_i, y_i)\,\Delta\beta_i \\ &+ b(t_i, y_i) b'(t_i, y_i) \int_{t_i}^{t_{i+1}} \int_{t_i}^{s} d\beta_u d\beta_s \end{aligned} \tag{2.46}$$

From Itô calculus it can be shown that the multiple stochastic integral

$$\int_{t_i}^{t_{i+1}} \int_{t_i}^{s} d\beta_u d\beta_s = \tfrac{1}{2}((\Delta\beta_i)^2 - \Delta t_i) \tag{2.47}$$

The Milshtein approximation (2.46) has a strong order of convergence $\lambda = 1.0$, see (Kloeden & Platen, 1992), which is higher than that of the Euler approximation. Analogous stochastic Taylor expansions also hold for the multidimensional case and for higher order of truncation, but they are much more complicated.

Higher order of strong and weak convergence can be obtained by including more terms from the stochastic Taylor expansion. Generally though, they are cumbersome to implement as numerical schemes because they involve increasingly higher order derivatives of the drift and diffusion coefficients of the stochastic differential equation. There are time discrete approximations which avoid the use of derivatives. These are called general *Runge-Kutta* type approximations for sde. Rümelin (1982) considers a very general explicit Runge-Kutta scheme. He shows that under certain boundary conditions the scheme converges uniformly in quadratic mean to the solution x_t of

$$dx_t = \left[a(t, \boldsymbol{x}_t) + \lambda\, b(t, \boldsymbol{x}_t)\frac{\partial b(t, \boldsymbol{x}_t)}{\partial x}\right] dt + b(t, x_t)\, d\beta_t \tag{2.48}$$

where the correction factor λ is a real number, defined through the specific scheme used. Eulers method is of Runge-Kutta type with $\lambda = 0$. Another commonly used integration scheme is *Heun's* method, which is also of Runge-Kutta type:

$$\begin{aligned} y_{i+1} &= y_i + \frac{1}{2}[a(t_i, y_i) + a(t_{i+1}, \tilde{y}_{i+1})]\, \Delta t_i \\ &+ \frac{1}{2}[b(t_i, y_i) + b(t_{i+1}, \tilde{y}_{i+1})]\, \Delta\beta_i \end{aligned} \tag{2.49}$$

with the Euler predictor

$$\tilde{y}_{i+1} = y_i + a(t_i, y_i)\, \Delta t_i + b(t_i, y_i)\, \Delta\beta_i \tag{2.50}$$

The quadratic mean convergence of this scheme is to the Itô solution of (2.48) with $\lambda = 1/2$, which corresponds to the solution of (2.14) interpreted in the sense of Stratonovich. It is shown by Rümelin (1982) that under certain conditions Heun's method has a global mean square error of order $O(h^2)$, where h is the stepsize. This is the same order as for the Milshtein scheme, whereas Eulers method only has a global error of order $O(h)$. It should be noticed that if one replaces a by $a - \lambda b b'$ in the given integration scheme one obtains convergence to the solution of the Itô equation (2.14).

2.6 Summary

In this Chapter it has been argued for stochastic differential equations being an obvious choice for modelling a wide range of physical systems. The usual way to model a number of physical systems using ordinary differential equations is extended by models of the noise of the systems. Usually it is necessary to include this noise model, to obtain a realistic representation of the system. The mathematical and computational treatment of the stochastic model is though more complicated than for the corresponding deterministic model.

It is shown how to simulate the sample paths from a stochastic differential equation. The basic element of the stochastic simulations is the pseudo random number generator. The verification of such generators is discussed in Appendix. The next to decide on is how to perform the stochastic integration. The most simple being the Euler integration. In order to obtain a higher accuracy in the simulation one can either increase the number of subintervals in the discretization of the time or resort to some higher order integration scheme, e.g. general Runge-Kutta schemes.

Chapter 3

Parameter Estimation

Assume that a certain model structure $\mathcal{M}$ has been selected. A set of candidate models is parameterized using the finite dimensional parameter vector $\theta \in \mathbf{\Theta}$. A particular model is denoted $\mathcal{M}(\theta)$. The problem is now to find the best model among the set $\mathcal{M}^*$

$$\mathcal{M}^* = \{\mathcal{M}(\theta) | \theta \in D_{\mathcal{M}}\} \tag{3.1}$$

where $D_{\mathcal{M}}$ is a closed subset of $\mathbb{R}^d$, where d is the dimension of the parameter vector.

The true system is denoted by $\mathcal{S}$, and the experimental condition under which the system operates by $\mathcal{X}$. This includes the properties of the input signal and possible feedback configurations. Sets of experimental conditions are denoted by X.

The definition of a *good* model is specified through its loss function. According to Ljung (1987) most of the methods used for parametric estimation can be characterized as general *prediction error methods*.

3.1 Prediction Error Methods

In the following it is assumed that a finite set of observations are obtained from the system $\mathcal{S}$, measured with equally spaced time intervals, with unit sampling time. Hence the time index $k \in \{0, 1, 2, \cdots, N\}$. In order to derive the loss function the following set of observations is introduced,

$$\mathbf{y}^k = [\mathbf{y}_k, \mathbf{y}_{k-1}, \cdots, \mathbf{y}_1, \mathbf{y}_0] \tag{3.2}$$

i.e. $\mathbf{y}^k$ is a matrix containing all the observations up to and including time t_k, and $\mathbf{y}_k \in \mathbb{R}^s$, where s is the dimension of the vector. The $\mathbf{y}$'s are considered to be the outputs from the model, i.e. dependent variables. Correspondingly the matrix of inputs to the model $\mathbf{u}^k$ is defined, where $\mathbf{u}_k \in \mathbb{R}^m$. The inputs may be controlled or not, but they are assumed to be measurable. Furthermore the inputs may be generated from either open- or closed-loop operation of the system. In closed-loop operation we assume that the input is generated as a feedback of the output according to

$$\mathbf{u}_k = \nu(t_k, \mathbf{y}^k, \mathbf{u}^{k-1}, \xi(t_k)) \tag{3.3}$$

where $\nu(\cdot, \cdot, \cdot, \cdot)$ is a given deterministic function and $\xi(t_k)$ is a deterministic signal, i.e. an external input. Closed-loop operation may be necessary for some systems, e.g. for unstable plants, but data obtained from such experiments can easily be defective, i.e. not informative enough, see e.g.

(Ljung, 1987). The joint set of the data (input and output) is defined as

$$z^k = (\mathbf{y}^k, \mathbf{u}^k) \tag{3.4}$$

The problem of estimation is a matter of how to use the information contained in the data z^N to select a proper value $\hat{\theta}_N$, and hence a proper model $\mathcal{M}(\hat{\theta}_N)$ from the set $\mathcal{M}^*$.

The performance of a model is judged by its ability to predict the outputs of a system. The one-step prediction error from a certain model $\mathcal{M}(\theta)$ is given by

$$\epsilon(t_k, \theta) \triangleq \mathbf{y}_k - \hat{\mathbf{y}}_k \tag{3.5}$$

$$\hat{\mathbf{y}}_k = \mathbf{g}(t_k, \theta, z^{k-1}) \tag{3.6}$$

where $\mathbf{g}(t_k, \theta, z^{k-1})$ is a deterministic function of old data and the parameters. If we are using off-line methods, the whole input sequence may be considered as known and (3.6) can be replaced by

$$\hat{\mathbf{y}}_k = \mathbf{g}(t_k, \theta, \mathbf{y}^{k-1}, \mathbf{u}^N) \tag{3.7}$$

A good model is now one that produces small prediction errors for a given data set, z^N. The idea is to choose some norm that measures the size of ϵ, and then find the parameter vector $\hat{\theta}_N$ that minimizes this norm. The criterion may be even further specialized as in the following. Let the prediction error sequence be filtered through a stable linear filter $\mathbf{L}(q)$,

$$\epsilon_f(t_k, \theta) = \mathbf{L}(q)\epsilon(t_k, \theta) \tag{3.8}$$

for all $k \in \{1, 2, \cdots, N\}$. Then use the following norm,

$$V(\theta, z^N) = \frac{1}{N} \sum_{k=1}^{N} l(t_k, \theta, \epsilon_f(t_k, \theta)) \tag{3.9}$$

where $l(\cdot)$ is a scalar valued function. The estimate $\hat{\theta}_N$ is chosen as the minimizing value of (3.9),

$$\hat{\theta}_N = \arg \min_{\theta \in D_{\mathcal{M}}} V(\theta, z^N) \tag{3.10}$$

Following Ljung (1987) methods that corresponds to the approach of (3.10) are called *prediction error methods*. This approach contains as special cases a number of known methods, like the *least squares* and *maximum likelihood* methods. The different methods apply by specific choices of the prefilter $L(q)$ and the norm $l(\cdot)$.

The effect of the filter L is easy to understand in a frequency domain interpretation of the criterion (3.10). L acts like a frequency weighting of the criterion, i.e. if L is a low pass filter then the criterion will only be little affected by high frequency disturbances. By another choice of the prefilter it is possible to remove slow drift terms in the data. It should be noted that the inclusion of the filter is to allow extra freedom in dealing with the properties of the prediction errors. The filter can be considered as a part of the model, by changing the predictor. If we consider a model of the form

$$y_k = G(q, \theta)u_k + H(q, \theta)e_k \tag{3.11}$$

then the effect of prefiltering the prediction errors according to (3.8) is identical to changing the noise model from $H(q, \theta)$ to

$$H_L(q, \theta) = L^{-1}(q)H(q, \theta) \tag{3.12}$$

see Ljung (1987).

In the current treatment of the parameter estimation, it is assumed that all the data z^N are available for estimating $\hat{\theta}_N$, this is known as off-line estimation. In some cases it may be interesting to have estimates $\hat{\theta}_k$ for all k, with the data set $\mathcal{Z}^k$, this is known as on-line estimation. One way to accomplish this is to make an off-line estimation at each sampling instant. This will include all available information up to time t_k, but the approach may involve a large number of calculations. Instead, a recursive algorithm of the following form is formulated

$$\begin{aligned} \mathbf{x}_k &= \mathbf{H}(t_k, \mathbf{x}_{k-1}, \mathbf{y}_k, \mathbf{u}_k) \\ \hat{\theta}_k &= \mathbf{G}(\mathbf{x}_k) \end{aligned} \tag{3.13}$$

where $\mathbf{x}_k$ is a state-vector of fixed dimension (typically $\dim(\mathbf{x}) < N$), that represents some information state. $\mathbf{H}$ and $\mathbf{G}$ are explicit expressions with a fixed amount of calculations. By applying these restrictions and formulating the recursive algorithm (3.13) for the parameter estimation it can be secured that the calculation of $\hat{\theta}_k$ can be evaluated during a sampling interval. The drawback in this approach is, that we do not make use off all available information in the calculation of $\hat{\theta}_k$. With an increasing amount of available computer power, the need for applying the restrictions of (3.13) for parameter estimation is vanishing. For the estimation cases in this report only off-line methods have been applied.

3.2 The Maximum Likelihood Method

We now turn to a special case of the prediction error method, which has been used in most of the cases in this report, the maximum likelihood method. It has been chosen for several reasons. Beyond the appealing concept of formal statistical inference, it has nice asymptotic properties under mild conditions. This makes it very useful for other aspects of identification, e.g. for model validation by using different kinds of likelihood based statistical tests. Such aspects are addressed in Chapter 5.

3.2.1 Principle of Likelihood

In statistical inference, the observations are considered as realizations of a stochastic variables. The observations from an experiment are represented by $\mathbf{y}^N$, see (3.2). For convenience the inputs, $\{\mathbf{u}^k\}$, are omitted in the following derivations. Since we are concerned with off-line methods it is assumed that the whole sequence of inputs are known a priori.

The likelihood function is the joint probability density of all the observations assuming that the parameters are known,

$$\begin{aligned}\mathcal{L}(\theta, \mathbf{y}^N) &= p(\mathbf{y}^N|\theta)\\ &= p(\mathbf{y}_N|\mathbf{y}^{N-1}, \theta)p(\mathbf{y}^{N-1}|\theta)\\ &= \prod_{k=1}^{N} p(\mathbf{y}_k|\mathbf{y}^{k-1}, \theta)\end{aligned} \tag{3.14}$$

where successive applications of the rule $p(a, b) = p(a \mid b)\, p(b)$ are used to express the likelihood function as a product of conditional densities.

It is now assumed that the sequence of innovations, $\{\epsilon_k\}$ are zero-mean, independent stochastic variables with the probability density function $p(\epsilon_k(\theta)|\theta)$. We are then able to develop expression (3.14) for the likelihood function. Now

$$\begin{aligned}p(\mathbf{y}_k|\mathbf{y}^{k-1}, \theta) &= p(\epsilon_k(\theta)|\mathbf{y}^{k-1}, \theta)\\ &= p(\epsilon_k(\theta)|\theta)\end{aligned} \tag{3.15}$$

and the following likelihood function is obtained

$$\mathcal{L}(\theta, \mathcal{Y}^N) = \prod_{k=1}^{N} p(\epsilon_k(\theta)|\theta) \tag{3.16}$$

The maximum likelihood estimate is found as the parameter vector that maximizes the likelihood function (3.16). Maximizing this function is the same as minimizing the function $-\log \mathcal{L}(\theta, \mathcal{Y}^N)$, which gives the expression for the estimator

$$\hat{\theta}_{ml} = \arg \min_{\theta \in D_{\mathcal{M}}} - \sum_{k=1}^{N} \log p(\epsilon_k(\theta)|\theta) \tag{3.17}$$

The maximum likelihood method can thus be seen to be a special case of the prediction error criterion (3.9) and (3.10). When the prediction errors are assumed to be Gaussian with zero mean, and covariance matrix $\mathbf{R}_k(\theta)$, we have

$$-\log \mathcal{L}(\theta, \mathcal{Y}^N) = \frac{1}{2} \sum_{k=1}^{N} \left(\epsilon_k^T \mathbf{R}_k^{-1} \epsilon_k + \log \det \mathbf{R}_k + s \log 2\pi \right) \tag{3.18}$$

where s is the dimension of $\mathcal{Y}$. The simple expression for the criterion arise from the fact, that the Gaussian distribution is characterized alone by their mean and covariance. The implementation of the algorithm and the use of a Newton algorithm for the optimization is denoted to Appendix A, see also (Madsen & Melgaard, 1991; Melgaard et al., 1992a; Melgaard & Madsen, 1993).

3.2.2 Separation of Filtering and Parameter Estimation

The estimation method requires access to the residuals, the difference between the predictions and the measurements. For the general class of models, the computation of the optimal predictions amounts to solving the general filtering problem. Exact solutions to this problem is only possible in special cases, e.g. for linear models, otherwise it is necessary to employ certain approximations to obtain implementable algorithms. However, it turns out, that while the properties of the ML parameter estimates depend upon the residuals and their properties, they do not depend on how they were found. It is therefore natural to separate the filtering problem from the parameter estimation problem. In that way we may develop the parameter estimator independently of the predictor, which according to the previous discussion, Ljung (1987), *is* the model itself. In the present derivation of the ML estimator there is only certain requirements on the residuals. As stated earlier the problem of treating a general class of models, is focused on the development of implementable predictors. This problem is treated in Chapter 4.

3.3 Asymptotic properties of parameter estimates

In this section the properties of $\hat{\theta}_N$ will be analyzed as N tends to infinity. The analysis is relevant for the confidence one should put on an estimate and hence a tool for comparing different estimators.

When discussing the asymptotic properties of an estimate, there will always be a tradeoff between the required conditions on the system and data, and the strength of statements about the properties. The usual assumption that the model set contains the true system, $\mathcal{S} \in \mathcal{M}$, for a given set of experimental conditions, $\mathcal{X}$, will often *not* be fulfilled for real systems. Usually the system is far more complex than we would allow the model to be.

For practical applications the objective of system identification is often to find an approximate description, catching the *relevant* features of the system. The discussion of consistency is therefore mainly of theoretical interest. On one hand if the identification method is not able to estimate the true system within a model set, it is probably not a good method. On the other hand, knowing that a method is consistent implies nothing a priori about its performance in *approximating* a more complex system, and the latter property is the most important when identifying a real process.

Even though we mainly consider the ML method, most of the results in this section apply to the more general prediction error method applied to the same model structure. This implies that if the conditions on the distribution of innovations required by the ML method cannot be obtained, we are still using a prediction error method. The asymptotic model thus obtained gives under very general conditions, the best average prediction performance. This result also implies a strong robustness property of the method (Ljung, 1978).

3.3.1 Convergence

Lets introduce some regularity assumptions from Ljung (1978) and Ljung & Caines (1979) which are relevant for analyzing the convergence properties of the parameter estimates.

The first condition is on the system. It requires that the expectation $E(y_k)$ exists and the system $\mathcal{S}$ is described by

$$y_k = E(y_k|\mathcal{Z}^{k-1}) + \epsilon_k \tag{3.19}$$

where the sequence of innovations, $\{\epsilon_k\}$ is a stochastic process with $E(\epsilon_k|\mathcal{Z}^{k-1}) = 0$. In 3.19 the conditional expectation is $E(y_k|\mathcal{Z}^{k-1}) = g_{\mathcal{S}}(t_k, y^{k-1}, u^{k-1})$, i.e. a deterministic function of old data.

A second condition is on the data, which concerns the system $\mathcal{S}$ and the experimental condition $\mathcal{X}$. The closed-loop system (3.3), (3.19) should be exponentially stable. That is, the past is forgotten at an exponential rate. If the system is linear, this condition simply requires that the poles of the system is inside the unit circle.

The next condition concerns the model set $\mathcal{M}$. It is required that $g(t_k, \theta, z^{k-1})$ in (3.6) is three times continuously differentiable with respect to θ. There is also a restriction on how fast g may increase with u_k and y_k for nonlinear models. Effectively, it may not increase faster than linearly. Another restriction is that the model and its derivatives with respect to θ are exponentially stable.

Consider a stochastic, linear state-space model given by the equations

$$x_{k+1} = A(t_k, \theta)x_k + B(t_k, \theta)u_k + E(t_k, \theta)v_k \tag{3.20}$$

$$y_k = C(t_k, \theta)x_k + D(t_k, \theta)u_k + F(t_k, \theta)e_k \tag{3.21}$$

where $v \sim N(0, I)$ and $e \sim N(0, I)$ are independent. Assume that the matrix elements are continuously differentiable with respect to $\theta \in D_{\mathcal{M}}$, where $D_{\mathcal{M}}$ is a compact set. If the system is completely *observable* and *controllable*, uniformly in t and in $\theta \in D_{\mathcal{M}}$, then the model fulfills the conditions on the model set previously described, and the predictor is the associated Kalman filter, see (Ljung, 1978).

The final conditions to set up, concerns the loss function $l(t_k, \theta, \epsilon)$ as given in the general prediction error formulation (3.9). It is required that the loss function is three times continuous differentiable with relation to θ and ϵ and that these differentials are bounded, (Ljung & Caines, 1979). In our case, where the maximum likelihood method is used, the loss function is given by the negative logarithm of the prediction error density, usually assumed to be Gaussian (3.17) and (3.18). Then the requirements on the derivatives in ϵ are fulfilled and the condition becomes mainly a question of parameterization (Graebe, 1990b).

When studying convergence of parameter estimates, we consider the general prediction error criterion

$$V(\theta, z^N) = \frac{1}{N} \sum_{k=1}^{N} l(t_k, \theta, \epsilon(t_k, \theta)) \tag{3.22}$$

which also contains the maximum likelihood method based on the Gaussian distribution. The limit

$$\bar{V}(\theta) = \lim_{N \to \infty} V(\theta, z^N) \tag{3.23}$$

exist under weak conditions on $\mathcal{S}$ and $\mathcal{X}$, e.g. if the processes are asymptotically stationary or periodic (Ljung & Caines, 1979). However this may not be valid for general time-varying and adaptive feedback, but even when the limit in (3.23) does not exist it is still possible to make statements about the convergence of the parameter estimates. If the limit exist the parameter estimate, $\hat{\theta}_N$ that minimizes the criterion converges to the set

$$D_{\bar{V}} = \{\theta | \theta \in D_{\mathcal{M}}; \bar{V} \leq \min_{\psi \in D_{\mathcal{M}}} \bar{V}(\psi)\} \tag{3.24}$$

If however the limit does not exist, we define the set

$$D_V = \{\theta | \theta \in D_{\mathcal{M}}; \liminf_{N\to\infty} E(V(\theta)) \leq \min_{\psi \in D_{\mathcal{M}}} \limsup_{N\to\infty} E(V(\psi))\} \tag{3.25}$$

This is clearly more general than (3.24) in that it does not require the existence of the limit $\bar{V}$, if however the limit exists then $D_V = D_{\bar{V}}$. We are now able to formulate the general theorem about parameter convergence.

Theorem 3.1 (Convergence) *Let the conditions previously described in this section be satisfied. Then*

$$\sup_{\theta \in D_{\mathcal{M}}} |V(\theta, z^N) - E(V(\theta, z^N))| \to 0 \quad \text{w.p.1 as } N \to \infty \tag{3.26}$$

uniformly in $\theta \in D_{\mathcal{M}}$. *Moreover, since the estimate* $\hat{\theta}_N$ *minimizes* $V(\theta, z^N)$, *it follows that*

$$\hat{\theta}_N \to D_V \quad \text{w.p.1 as } N \to \infty \tag{3.27}$$

if the limit (3.24) exists then $D_V = D_{\bar{V}}$. *For a proof see (Ljung, 1978).*
□

The theorem states that the value of the loss function calculated from a realization of data will become arbitrarily close to its expected value as the data length approaches infinity, and that parameter estimates computed by minimizing the loss function will converge into appropriate sets. Note that in the conditions of the theorem it is not required that the model set contains a system equivalent member.

3.3.2 Distribution of Parameters

Having established the conditions for convergence of the parameter estimates it will now be established that the parameter estimates are asymptotically normally distributed under similar conditions. Also in this case the approximate modelling approach of Ljung & Caines (1979) is adopted.

Theorem 3.2 (Distribution) *Assume that the conditions of theorem 3.1 are satisfied. Consider the general loss function (3.9)*

$$V(\theta, z^N) = \frac{1}{N} \sum_{k=1}^{N} l(t_k, \theta, \epsilon(t_k, \theta)) \tag{3.28}$$

Let $\hat{\theta}_N$ *be the global minimum of (3.28) over the compact set* $D_{\mathcal{M}}$*. Introduce the function* $W_N(\theta) = E(V(\theta, z^N))$*, where the expectation is with respect to the data. Let* θ_N^* *be the global minimum of* $W_N(\theta)$*. Now introduce the term*

$$P_N = (W_N''(\theta_N^*))^{-1} U_N (W_N''(\theta_N^*))^{-1} \tag{3.29}$$

where a quote denotes differentiation w.r.t. the parameters, and

$$U_N = E(NV'(\theta_N^*, z^N)(V'(\theta_N^*, z^N))^T) \tag{3.30}$$

and assume that $W_N''(\theta_N^*)$ *and* U_N *are invertible. Then the quantity* $\sqrt{N} P_N^{-1/2}(\hat{\theta}_N - \theta_N^*)$ *is asymptotically normal with zero mean and unit covariance matrix. Assume furthermore that*

$$W_N(\theta) \rightarrow \bar{W}(\theta) \tag{3.31}$$

uniformly in $D_{\mathcal{M}}$ *as* $N \rightarrow \infty$*. Let* $\bar{W}(\theta)$ *have a global minimum at* θ^**, and assume that* $\bar{W}''(\theta^*)$ *is invertible. Let* U_N *be defined as (3.30),* *and assume the limit* $U = \lim_{N\to\infty} U_N$ *exists and is invertible. If now*

$$\sqrt{N}W'_N(\theta^*) \to 0 \quad \text{as } N \to \infty, \tag{3.32}$$

then $\sqrt{N}(\hat{\theta}_N - \theta^*)$ *is asymptotically normal with zero mean and covariance matrix*

$$P = (\bar{W}''(\theta^*))^{-1} U (\bar{W}''(\theta^*))^{-1} \tag{3.33}$$

□

The theorem is an adaption of theorem 1 and the corollary in Ljung & Caines (1979), and a proof can be found there. The basic conditions are the same as required for the convergence theorem, but the requirement of a unique global minimum of the criterion function is fairly restrictive. From theorem 3.1 we know that the minimization of V will lead us close to a local minimum of W. We may then think of this theorem as applying to the neighborhood of this local minimum point. The assumption, in the second part of the theorem, of convergence of $W_N(\theta)$ as N goes to infinity is satisfied, for example, if the processes are asymptotically stationary or periodic.

If the true parameter θ_0 is contained in the model set such that $\{\epsilon(t_k, \theta_0)\}$ is a sequence of independent random vectors, results from the asymptotic distribution of the deviation of the estimate from θ_0 can be used to obtain confidence intervals for the parameters etc. If the limiting model θ^* does not give a true description of the system (i.e., if $\{\epsilon(t, \theta^*)\}$ are not independent) but only the best approximation available in $\mathcal{M}$, then the distribution may still be used for model validation, concerning e.g. the relevance of certain parameters in the model, see (Ljung & van Overbeek, 1978).

Cramér-Rao Bound

As a quality measure of an estimator we use its mean square error matrix

$$P(\hat{\theta}_N) = E(\hat{\theta}_N - \theta_0)(\hat{\theta}_N - \theta_0)^T \tag{3.34}$$

where θ_0 denote the true parameter. We are interested in selecting an estimator that makes P small. There is a lower limit to the values of P that can be obtained by any unbiased estimator. This is shown by the following.

Theorem 3.3 (Cramér-Rao inequality) *Let* $\hat{\theta}_N$ *be an estimator of* $\theta \in D_{\mathcal{M}}$*, where* $D_{\mathcal{M}}$ *is a compact subset in* $\mathbb{R}^d$ *such that* $E(\hat{\theta}_N) = \theta_0$*. The expectation is with relation to the data, assuming that the probability density function of the data is* $p(\mathbf{y}^N|\theta_0)$*, for all values of* θ_0*. Suppose that* $\mathbf{y}^N$ *may take values in a subset of* $\mathbb{R}^{N\times s}$*, whose boundary does not depend on* θ*. Then under mild regularity conditions*

$$P(\hat{\theta}_N) \geq \mathbf{M}^{-1} \tag{3.35}$$

where

$$\mathbf{M} \triangleq E(\tfrac{\partial}{\partial\theta} \log p(\mathbf{y}^N|\theta)(\tfrac{\partial}{\partial\theta} \log p(\mathbf{y}^N|\theta))^T)|_{\theta=\theta_0} \tag{3.36}$$

is known as the Fisher information matrix*.*

PROOF: *see e.g. Goodwin & Payne (1977).* □

Note that the evaluation of M requires knowledge of θ_0, so the exact value of M is usually not available to the user. If the covariance of an unbiased estimator achieves the lower bound of the Cramér-Rao inequality, it is said to be *efficient*.

We now return to the ML estimator (3.17), assuming independent innovations with a known probability density function. Assuming that $\mathcal{S} \in \mathcal{M}$ this estimator converges to a normal random variable according to theorem 3.2, with a covariance matrix $\mathbf{P}(\hat{\theta}) = \mathbf{M}^{-1}$, i.e. it attains, asymptotically, the lower limit of the Cramér-Rao inequality, see Goodwin & Payne (1977). Another interesting property about efficient estimators is that whenever there exists an unbiased efficient estimator, then it is also the maximum likelihood estimator. This is shown e.g. by Goodwin & Payne (1977).

3.3.3 Consistency

Suppose, that the model set *contains* a system equivalent member, corresponding to a parameter value θ_0, does then the estimate $\hat{\theta}_N$ tend to this *true* value of the parameters as the number of data tends to infinity? This matter of consistency is discussed in this section.

A suitable way to express that the true system is representable within the model set $\mathcal{M}$ is to require that the set

$$\begin{aligned} D_T(\mathcal{S}, \mathcal{M}, X) = \{&\theta | \theta \in D_{\mathcal{M}}; \\ &\lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} E|\hat{\mathbf{y}}_{\mathcal{S}}(t_k) - \hat{\mathbf{y}}_{\mathcal{M}}(t_k|\theta)|^2 = 0, \forall \mathcal{X} \in X\} \end{aligned} \tag{3.37}$$

is nonempty, where X is the set of experimental conditions, under which we would like our model to be valid. Here $\hat{\mathbf{y}}_{\mathcal{S}}(t_k) = E(\mathbf{y}_k|\mathbf{y}^{k-1})$ assuming condition 3.19 is satisfied, and $\hat{\mathbf{y}}_{\mathcal{M}}(t_k|\theta)$ is the predictor as given in (3.6). Hence if D_T is nonempty and $\theta^* \in D_{\mathcal{M}}$, then the corresponding model $\mathcal{M}(\theta^*)$ is, in the mean square sense of predicted output, indistinguishable from the true system, (Ljung, 1978).

Theorem 3.4 (Consistency) *Assume that the conditions of Theorem 3.1 apply. Assume also the following condition on the criterion functions*

$$\theta_0 \in D_T(\mathcal{S},\mathcal{M},X) \Rightarrow \forall k, \theta \in D_{\mathcal{M}} : El(t_k,\theta_0,\epsilon) \leq El(t_k,\theta,\epsilon) \tag{3.38}$$

The condition is satisfied e.g. for the ML criterion (3.18). Assume that the parameter set of mean square equivalence, $D_T(\mathcal{S},\mathcal{M},X)$*, defined in (3.37) is nonempty. For a particular experiment,* $\mathcal{X} \in X$*, we define the set*

$$D_I(\mathcal{S},\mathcal{M},\mathcal{X}) =$$
$$\{\theta | \theta \in D_{\mathcal{M}}; \liminf_{N\to\infty} \frac{1}{N}\sum_{k=1}^{N} E|\hat{y}_{\mathcal{S}}(t_k) - \hat{y}_{\mathcal{M}}(t_k|\theta)|^2 = 0\} \tag{3.39}$$

Then

$$\hat{\theta}_N \to D_I(\mathcal{S},\mathcal{M},\mathcal{X}) \quad \text{w.p.1 as } N\to\infty \tag{3.40}$$

□

The theorem is an adaption of lemma 4.1 in Ljung (1978). The basic conditions of the theorem are the same as required for the asymptotic convergence and distribution theorems, but additionally it is required that D_T from (3.37) is nonempty. This is a strong condition, but the discussion of consistency is only meaningful if the model set contains a system equivalent member in some sense. The set D_T represents the *correct* descriptions of the system in the sense of mean square equivalence. It is obvious that

$$\forall \mathcal{X} \in X : D_T(\mathcal{S},\mathcal{M},X) \subseteq D_I(\mathcal{S},\mathcal{M},\mathcal{X}) \tag{3.41}$$

since D_I is dependent of a certain experimental condition $\mathcal{X}$. If, under this experimental condition, we are not able to get a sufficiently representative picture of the system, then we are not guaranteed convergence to the right model, according the theorem. The desired relation is that

$$D_T(\mathcal{S}, \mathcal{M}, X) = D_I(\mathcal{S}, \mathcal{M}, \mathcal{X}) \tag{3.42}$$

for the chosen experimental condition $\mathcal{X}$. This is a matter of choosing the right $\mathcal{X}$, which includes properties like *persistently exciting* inputs, and not too special feedback mechanisms. This discussion of experimental design is postponed to Chapter 6. Even when (3.42) is satisfied, and theorem 3.4 guarantees the convergence to a system equivalent model, there might still be several parameter values yielding this property. In order to obtain uniqueness of the solution further restrictions on the model structure and parameterization must be applied. This is a matter of structural identifiability which is considered in Section 6.1.

3.4 Maximum Aposteriori Estimate

Compared to classical estimation, the Bayesian approach gives a conceptually different treatment of the parameter estimation problem. In the Bayesian approach the parameter itself is considered as a random variable with some prior probability density

$$\theta \sim p(\theta) \tag{3.43}$$

Then based on observations of other random variables, the data set $\mathbf{y}^N = [\mathbf{y}_N, \mathbf{y}_{N-1}, \cdots, \mathbf{y}_1, \mathbf{y}_0]$, which is correlated with the parameters, we may infer information about its value. Suppose the conditional probability density of the observations, given θ, is

$$p(\mathbf{y}^N|\theta), \tag{3.44}$$

which is the likelihood function. Then by combining the prior information with the sample information, using Bayes' theorem, it is possible to form the posterior information for θ, i.e. the conditional probability density given the observations

$$p(\theta|\mathbf{y}^N) = \frac{p(\mathbf{y}^N|\theta)p(\theta)}{p(\mathbf{y}^N)} \propto p(\mathbf{y}^N|\theta)p(\theta) \tag{3.45}$$

It is seen, that the posterior probability density is proportional to the likelihood function multiplied by the prior probability density. From the posterior probability density, different point estimates of θ can be determined, for instance the value where the probability density attains its maximum. This is called the maximum a posteriori estimate (MAP),

$$\hat{\theta}_{map} = \arg\max_{\theta \in \Omega} p(\mathbf{y}^N|\theta)p(\theta) \tag{3.46}$$

over an admissible parameter set Ω. The conventional maximum likelihood estimator (ML) selects the parameter that maximizes the likelihood of the data,

$$\hat{\theta}_{ml} = \arg\max_{\theta \in \Psi} p(\mathbf{y}^N|\theta) \tag{3.47}$$

If the prior distribution in 3.46 is uniform then the MAP estimator is equivalent to an ML estimator that is restricted to $\Psi = \Omega$. So the ML estimator corresponds to the MAP estimator when the prior information is diffuse or non-informative, with the constraint $\Psi = \Omega$.

Example 3.1 *Assume that data can be described by:*

$$\mathbf{y}_k = E(\mathbf{y}_k|\mathbf{y}^{k-1}, \theta) + \epsilon_k(\theta) \tag{3.48}$$

where the sequence of innovations $\{\epsilon_k(\theta)\}$ *is a stochastic process with the property* $E(\epsilon_k(\theta)|\mathbf{y}^{k-1}) = 0$*. If we assume the innovations are mu-*

tually independent and have the probability density function $p(\epsilon_k(\theta)|\theta)$, *we obtain the following likelihood function of the data:*

$$p(\mathbf{y}^N|\theta) = \prod_{k=1}^{N} p(\epsilon_k(\theta)|\theta) \tag{3.49}$$

If the innovations are assumed to be Gaussian with zero mean, and covariance $\mathbf{R}_k(\theta)$, *we have*

$$p(\mathbf{y}^N|\theta) \propto \prod_{k=1}^{N} (\det \mathbf{R}_k(\theta))^{-1/2} \exp(-\frac{1}{2}\epsilon_k(\theta)^T \mathbf{R}_k^{-1}(\theta)\epsilon_k(\theta)) \tag{3.50}$$

The maximizing argument of (3.50) with relation to θ *thus yields the ML estimate,* $\hat{\theta}_{ml}$. *Let the prior distribution function for* θ *be given by a normal distribution with mean* $\mu = \mu_\theta$ *and covariance* $\mathbf{\Sigma} = \mathbf{\Sigma}_\theta$. *Then*

$$p(\theta) \propto \exp(-\frac{1}{2}(\theta - \mu_\theta)^T \mathbf{\Sigma}_\theta^{-1}(\theta - \mu_\theta)) \tag{3.51}$$

Following (3.45), the conditional pdf for θ *given the data* $\mathbf{y}^N$, *that is the posterior pdf for* θ, *is given by*

$$\begin{aligned} p(\theta|\mathbf{y}^N) \propto {} & \exp(-\frac{1}{2}(\theta - \mu_\theta)^T \mathbf{\Sigma}_\theta^{-1}(\theta - \mu_\theta)) \times \\ & \prod_{k=1}^{N} (\det \mathbf{R}_k(\theta))^{-1/2} \exp(-\frac{1}{2}\epsilon_k(\theta)^T \mathbf{R}_k^{-1}(\theta)\epsilon_k(\theta)) \end{aligned} \tag{3.52}$$

The argument that maximizes (3.52) is the MAP estimate, $\hat{\theta}_{map}$. *The difference between ML and MAP is clearly seen from (3.50) and (3.52). As the prior standard deviation increases,* $\mathbf{\Sigma}_\theta \to \infty$, *the prior pdf approaches a uniform and non-informative distribution, and the MAP estimate becomes equivalent with the ML estimate.* □

The advantage of considering MAP estimators instead of ML estimators is of course that our prior information about the parameters of the model is incorporated in the final estimate of the parameters. We may then be able to reach the same level of confidence of the final estimates, from a shorter experiment. Or non-identifiable models may become identifiable, due to the external information.

The problematic part of the MAP estimation approach, is the specification of the prior pdf. The prior information can either be generated from samples of past data, e.g. from ML estimation, in which case the asymptotic distribution of the estimates is known. The prior information can also be specified from introspection, or theoretical considerations, and in this case the prior information may differ for individual persons, because they have different beliefs. One should also consider the sensibility of the final estimate from the specification of a wrong prior pdf. Often the prior pdf represents both kinds of prior information, data-based and non data-based.

The Bayesian approach is quite appealing in relation to grey box identification (Section 1.1), where the idea is to utilize the prior information about the system, in the identification. The MAP estimator is also called the grey-box estimator by Tulleken (1993).

It can be shown that asymptotically (when the number of data observations $N \to \infty$) the MAP estimator converge with probability one to the conventional ML estimator, provided the admissible model set contains the the true parameters. That is, the MAP estimator is asymptotically unbiased and efficient then the a priori knowledge is correct. This is a direct consequence of the fact that the likelihood will asymptotically become Gaussian distributed, around the true parameters, with vanishing covariance matrix. This also means that the influence from the prior distribution is vanishing with increasing number of observations in the likelihood function, see Tulleken (1993).

3.5 Robust Norms

From the previous sections it has been shown that for minimizing the variance of the estimated parameters the optimal choice of l for the general PEM criterion (3.9) is

$$l(\epsilon_t(\theta)) = -\log p(\epsilon_t(\theta)) \tag{3.53}$$

where $p(\epsilon_t(\theta))$ is the probability density function of the innovations (the ML criterion). The problem is that $p(\epsilon_t(\theta))$ may not be known for the true innovations. If we assume the density is Gaussian, but the true density has thicker tails, this may have a large influence on the estimation, cf. (Martin & Yohai, 1985). It is convenient to consider the innovations as being generated from an outliers model, with the following distribution of the innovations

$$P(\epsilon_t) = (1-\gamma)N(0,\sigma^2) + \gamma G\ , \qquad \gamma > 0 \tag{3.54}$$

where G is some symmetric outlier generating distribution and γ is a small fraction.

In order to evaluate the influence of the outliers on the estimates when using a given norm l, it is valuable to rewrite the expression for the covariance matrix of the parameters (3.33). Under the assumption that $\mathcal{S} \in \mathcal{M}$ and that the limiting parameters equals the true ones, $\theta^* = \theta_0$, the expression for the covariance of the parameters can be written,

$$\mathbf{P}_\theta = \kappa(l)\,(\bar{W}''(\theta_0))^{-1} \tag{3.55}$$

where the covariance of the parameters is scaled by the scalar

$$\kappa(l) = E_\epsilon \psi^2(\epsilon_t(\theta_0)) \;/\; (E_\epsilon \psi'_\epsilon(\epsilon_t(\theta_0)))^2 \tag{3.56}$$

where $\psi(\epsilon_t) = l'_\epsilon(\epsilon_t)$ (derivation is with respect to ϵ) and expectation should be taken with respect to the true distribution of the innovations. Consider for simplicity the scalar case of Gaussian ML with the given variance 1, then (3.53) is $l(\epsilon_t) = \frac{1}{2}\epsilon_t^2$ (neglecting the constant term). For the ideal case where the true distribution of the innovations is Gaussian we have $\kappa(l) = 1$. If on the other hand the true distribution is given by (3.54) with a $\gamma > 0$ this scalar can become much larger than 1, cf. (Ljung, 1987, pp. 397–398). This means, that even if the fraction of outliers is very small the variance of the parameters can be much deteriorated when using the norm (3.53). Thus, the norm (3.53) is very sensitive to the true distribution of the innovations. This is of course not desirable because the true distribution is usually unknown. In order to make the criterion more *robust* to the unknown variations of the distribution of the innovations, it is suggested to modify (3.53) in such a way that $\psi(x)$ in (3.56) behaves like x for small values of x, then saturates, and even tend to zero as x increases (redescending). One choice of a ψ-function, is called *Huber's psi*

$$\psi_H(x) = \begin{cases} x & |x| \le c_H \\ c_H \operatorname{sgn}(x) & |x| > c_H \end{cases} \tag{3.57}$$

If one wants protection against extremely heavy-tailed distributions, a redescending psi-function can be used. A frequently used type of redescending function is *Tukey's bisquare* function

$$\psi_B(x) = \begin{cases} x\,(1 - x^2/c_B^2)^2 & |x| \le c_B \\ 0 & |x| > c_B \end{cases} \tag{3.58}$$

see Martin & Yohai (1985) for a discussion of the psi-functions and robustness. When off-line methods are of greatest interest we have the possibility

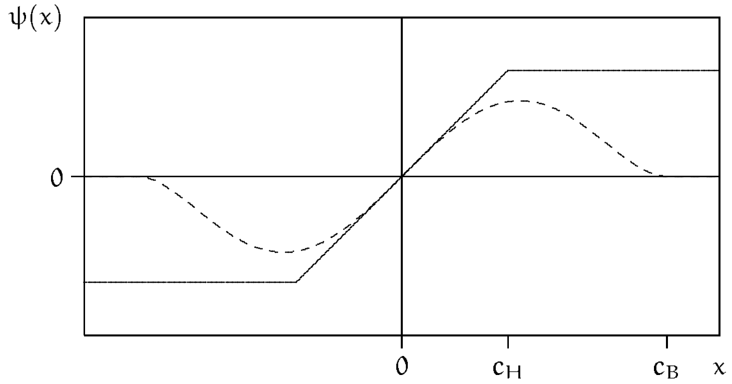


Figure 3.1. *Huber's psi-function and Tukey's bisquare function.*

of repeating the estimation on the data set. Hence, we are able to detect and handle outliers in the data set, e.g. in connection with the model validation, see Chapter 5. Thus for off-line methods it may not be necessary to be protected against very heavy-tailed distributions, but it is still useful to modify the estimation criterion according to e.g. Huber's psi-function and with a c_H value that is not too small, e.g. $c_H = 3\,\hat{\sigma}$. When such a robust norm is used there will only be a small increase of $\kappa(l)$ when outliers are present and also when the true distribution is Gaussian (no outliers) there is only a small loss of optimality compared to the original norm (3.53). The price of a small increase of the variance for the nominal case is thus worth paying to have robustness against small variations in the true distribution of the innovations.

3.6 Preprocessing of Data

When the data has been collected from an identification experiment it usually needs some preprocessing before it can be used by an estimation algorithm. There may be outliers in the data or missing observations in the data set. When redundant information is available, that is, more sensors measuring the same quantity of the system for a given simplified model, it must be decided how this information is compounded into a representative measurement. Preprocessing of the data can turn out to be a non-trivial task. But proper handling is required in this phase, because the resulting estimate is based on these data.

How redundant information is used to create a single representative measurement of a certain quantity, if this is required for a simplified modeling, is highly dependent of the individual case. Usually some weighted sum of the signals can be used. The weights can either be found by a principal components analysis as the weights for the first principal component or by physical arguments, e.g. that a temperature sensor represents a certain area of the surface to be measured. Ideally, if the redundant sensors have been placed properly, the first principal component will put equal weights on each sensor. If this is not the case, this could be a way to detect a failing sensor.

If outliers or missing values are detected in the data, e.g. by plotting the data or from residual analysis, there are different possibilities of handling the problem. To some degree the estimation algorithm can be protected against outliers by using robust methods as discussed in Section 3.5. Some estimation methods can be used regardless of missing observations in the outputs of the system. When a prediction error method is used, this is done e.g. by replacing the missing output with the predicted output.

When using off-line methods it is possible to repair the data before using it for model identification. If redundant or highly correlated signals are available, regression techniques can be used to replace or fill gaps of one signal based on other signals. Otherwise, robust dynamical methods can be used to estimate e.g. robust AR models of the timeseries followed by a robust smoother to replace the outliers or fill the gaps of missing values, see e.g. (Martin & Yohai, 1985).

3.7 Summary

In this chapter the general concept of parametric estimation of dynamical systems has been set up. The procedures deals with the sequence of prediction errors $\{\epsilon(t_k, \theta)\}$ computed from the proposed models using the observed data. In this sense we have separated the predictor (model) from the estimation procedure. It is, for example, of none importance for the estimation procedure if the model has a discrete time or continuous time formulation.

A number of asymptotic properties for the estimation procedures has been discussed. Results for convergence and distribution of the asymptotic estimate is shown for the general case, when $\mathcal{S} \notin \mathcal{M}$, which is called approximate modeling. In the case where the true parameter is available in the model set, results concerning efficiency and consistency have been presented.

There has mainly been focused on the maximum likelihood method, because of its attractive property as an asymptotically efficient estimator.

The MAP estimator has been presented as obvious choice for grey box identification by including the a priori distribution of the parameters in the estimation. The asymptotic properties of the MAP estimator are under reasonable conditions equal to those of the ML estimator. In all cases, though, it may be necessary to use a modification of the estimation criterion to obtain robustness against outliers in the data.

Chapter 4

Filtering the State

Referring to Chapter 2, let the model be described by the vector Itô stochastic differential equation

$$d\mathbf{x}_t = \mathbf{f}(\mathbf{x}_t, \mathbf{u}_t, \theta, t)dt + \mathbf{G}(\mathbf{x}_t, \theta, t)d\beta_t \tag{4.1}$$

with β being a standard Wiener process. The observations $\mathbf{y}_k$ are taken at discrete time instants, t_k

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k, \mathbf{u}_k, \theta, t_k) + \mathbf{e}_k \tag{4.2}$$

where $\mathbf{e}$ is a Gaussian white noise process independent of β, and $\mathbf{e}_k \sim N(0, \mathbf{S}(\theta, t_k))$. (The super- and subscript k is shorthand notation for t_k throughout the chapter). In order to estimate the parameters of the continuous time model (4.1) and (4.2), based on discrete time observations, we need access to the one-step prediction errors, defined in (3.5) and (3.7), repeated here

$$\epsilon_k(\theta) \triangleq \mathbf{y}_k - \hat{\mathbf{y}}_k \tag{4.3}$$

$$\hat{\mathbf{y}}_k = \mathbf{h}(\hat{\mathbf{x}}_k, \mathbf{u}_k, \theta, t_k) \tag{4.4}$$

$$= \mathbf{g}(t_k, \theta, \mathbf{y}^{k-1}, \mathbf{u}^N) \tag{4.5}$$

where the predictor, $\mathbf{g}(t_k, \theta, \mathbf{y}^{k-1}, \mathbf{u}^N)$ is a deterministic function of old outputs, the parameters and the inputs. In the following derivations the input sequence $\mathbf{u}^N$ is skipped for convenience, but since we are considering off-line methods the inputs are assumed to be known a priori. The issue now, is to formulate the predictor corresponding to the model (4.1) and (4.2). For the Gaussian maximum likelihood method (3.18), in addition to the one-step predictions, we also need the covariance of the predictions, for evaluation of the likelihood function. For non-Gaussian maximum likelihood we also need higher order moments or even the whole distribution $p(\mathbf{y}_k|\mathbf{y}^{k-1}, \theta)$.

To solve these problems, we must establish the conditional density for $\mathbf{x}_k$ conditioned on the measurements up to an including time t_k, $p(\mathbf{x}_k|\mathbf{y}^k)$. If this objective can be accomplished, then various estimators can be defined, optimal with respect to some specified criterion, such as the conditional mean or the conditional mode.

4.1 Exact Filtering

We will now describe the solution to the continuous-discrete filtering problem conceptually. We are considering a continuous time model with discrete time observations, therefore the term continuous-discrete filtering. We solve the problem conceptually, because only in certain special cases it is possible to implement the exact solution.

4.1.1 Conditional Density

Before establishing the results about the conditional density, we have to impose certain assumptions on the model (4.1).

Condition 4.1 (Itô Equation) *Suppose the real functions* $\mathbf{f}$ *and* $\mathbf{G}$*, and initial condition* $\mathbf{x}_0$*, satisfy the following conditions.* $\mathbf{f}$ *and* $\mathbf{G}$ *satisfy uniform Lipschitz conditions in* $\mathbf{x}$*:*

$$\begin{aligned}||\mathbf{f}(\mathbf{x}_2, t) - \mathbf{f}(\mathbf{x}_1, t)|| &\leq K||\mathbf{x}_2 - \mathbf{x}_1||,\\ ||\mathbf{G}(\mathbf{x}_2, t) - \mathbf{G}(\mathbf{x}_1, t)|| &\leq K||\mathbf{x}_2 - \mathbf{x}_1||.\end{aligned}$$

$\mathbf{f}$ *and* $\mathbf{G}$ *satisfy Lipschitz conditions in* t *on* $[t_0, T]$*:*

$$\begin{aligned}||\mathbf{f}(\mathbf{x}, t_2) - \mathbf{f}(\mathbf{x}, t_1)|| &\leq K||t_2 - t_1||,\\ ||\mathbf{G}(\mathbf{x}, t_2) - \mathbf{G}(\mathbf{x}, t_1)|| &\leq K||t_2 - t_1||.\end{aligned}$$

The initial condition $\mathbf{x}_0$ *is a random variable with* $E(||\mathbf{x}_0||^2) < \infty$*, independent of* $\{\beta_t, t \in [t_0, T]\}$*.* $\square$

If condition 4.1 is satisfied then $\{\mathbf{x}_t\}$ is a Markov process, and, in the mean square sense, is uniquely determined by the initial condition $\mathbf{x}_0$, see (Jazwinski, 1970).

The evolution of the probability density $p(\mathbf{x}_t|\mathbf{y}^k)$, $t \in [t_k, t_{k+1}[$ of the Markov process generated by the Itô equation (4.1) is described by a partial differential equation, which is known as *Kolmogorov's forward equation* or the *Fokker-Planck* equation.

$$dp(\mathbf{x}_t|\mathbf{y}^k) = \mathcal{L}(p)\, dt \qquad t \in [t_k, t_{k+1}[\tag{4.6}$$

where

$$\mathcal{L}(\,\cdot\,) = -\sum_{i=1}^{n} \frac{\partial(\,\cdot\, f_i)}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2(\,\cdot\,(\mathbf{G}\mathbf{G}^T)_{ij})}{\partial x_i x_j} \tag{4.7}$$

is the forward diffusion operator, see (Jazwinski, 1970; Maybeck, 1982). This is the evolution *between* observations. The initial condition, at t_k, is $p(\boldsymbol{x}_k|\boldsymbol{y}^k)$. We assume this density exists and is once continuously differentiable with respect to t and twice with respect to $\boldsymbol{x}$. It remains to determine how p changes at an observation at t_k, that is to determine the relationship between $p(\boldsymbol{x}_k|\boldsymbol{y}^k)$ and $p(\boldsymbol{x}_k|\boldsymbol{y}^{k-1})$. In addition to the earlier stated conditions we also assume that the observation mapping $\mathbf{h}$, given in (4.2), is continuous in $\boldsymbol{x}$ and in t and bounded for each t_k w.p.1. Since $p(\boldsymbol{x}_k|\boldsymbol{y}^k) = p(\boldsymbol{x}_k|\boldsymbol{y}_k, \boldsymbol{y}^{k-1})$ we can use Bayes' rule to obtain

$$p(\boldsymbol{x}_k|\boldsymbol{y}^k) = \frac{p(\boldsymbol{y}_k|\boldsymbol{x}_k, \boldsymbol{y}^{k-1})p(\boldsymbol{x}_k|\boldsymbol{y}^{k-1})}{p(\boldsymbol{y}_k|\boldsymbol{y}^{k-1})} \tag{4.8}$$

This expression can be simplified, since the process $\{\boldsymbol{e}_k\}$ is assumed to be Gaussian white noise, with $\boldsymbol{e}_k \sim N(0, \mathbf{S}_k)$,

$$\begin{aligned} p(\boldsymbol{y}_k|\boldsymbol{x}_k, \boldsymbol{y}^{k-1}) &= p(\boldsymbol{y}_k|\boldsymbol{x}_k) \\ &= ((2\pi)^s \det \mathbf{S}_k)^{-1/2} \exp(-\tfrac{1}{2}\tilde{\boldsymbol{y}}_k^T \mathbf{S}_k^{-1} \tilde{\boldsymbol{y}}_k) \end{aligned} \tag{4.9}$$

where $\tilde{\boldsymbol{y}}_k = \boldsymbol{y}_k - \mathbf{h}(\boldsymbol{x}_k, \boldsymbol{u}_k, \theta, t_k)$. Similarly, we can compute

$$p(\boldsymbol{y}_k|\boldsymbol{y}^{k-1}) = \int_{[x_k]} p(\boldsymbol{y}_k|\xi)p(\xi|\boldsymbol{y}^{k-1})\, d\xi \tag{4.10}$$

Hence equation (4.8) becomes

$$p(\boldsymbol{x}_k|\boldsymbol{y}^k) = \frac{p(\boldsymbol{y}_k|\boldsymbol{x}_k)p(\boldsymbol{x}_k|\boldsymbol{y}^{k-1})}{\int_{[x_k]} p(\boldsymbol{y}_k|\xi)p(\xi|\boldsymbol{y}^{k-1})\, d\xi} \tag{4.11}$$

with $p(\boldsymbol{y}_k|\boldsymbol{x}_k)$ given by (4.9).

Conceptually we now have the entire density function $p(\mathbf{x}_t|\mathbf{y}^k)$, $t \in [t_k, t_{k+1}[$ for all $k > 0$ for some given initial state $p(\mathbf{x}_0)$. Kolmogorov's forward equation describes the evolution between the observations and equation (4.11) gives the update formula, when a new observation is available at time t_k. The conditions we have imposed on the system are merely to guarantee the existence of the conditional probability density function.

With this density function available we are now able to calculate the terms $\hat{\mathbf{y}}_k$ and $\mathbf{R}_k$ required for the parameter estimation, using Gaussian maximum likelihood

$$\hat{\mathbf{y}}_k = \int_{[x_k]} \mathbf{h}(\xi, \mathbf{u}_k, \theta, t_k) p(\xi|\mathbf{y}^{k-1})\, d\xi \tag{4.12}$$

$$\begin{aligned}\mathbf{R}_k = &\int_{[x_k]} \mathbf{h}(\xi, \mathbf{u}_k, \theta, t_k)\mathbf{h}(\xi, \mathbf{u}_k, \theta, t_k)^T p(\xi|\mathbf{y}^{k-1})\, d\xi \\ &+\mathbf{S}_k - \hat{\mathbf{y}}_k \hat{\mathbf{y}}_k^T\end{aligned} \tag{4.13}$$

Thereby we have defined all the elements for conceptually solving the identification problem. The difficulty is that the solution requires the entire density function $p(\mathbf{x}_t|\mathbf{y}^k)$, $t \in [t_k, t_{k+1}[$, involving partial differential equations, which can not be solved exactly for general nonlinear models.

When $\mathbf{h}(\mathbf{x}_k, \mathbf{u}_k, \theta, t_k)$ is linear with respect to $\mathbf{x}_k$ the situation is less complicated. Then, when using Gaussian ML, the distribution of $p(\mathbf{x}_t|\mathbf{y}^k)$, $t \in [t_k, t_{k+1}[$ is also Gaussian and only the first two moments of the distribution is needed.

4.1.2 Evolution of Moments

Let $\varphi(\mathbf{x})$ be a twice continuously differentiable scalar function of the state-vector $\mathbf{x}$. Define

$$E^k(\varphi(\mathbf{x}_t)) \triangleq E(\varphi(\mathbf{x}_t)|\mathbf{y}^k) = \int_{[x_t]} \varphi(\xi)p(\xi|\mathbf{y}^k)\,d\xi \tag{4.14}$$

the expectation of φ using the conditional density $p(\mathbf{x}_t|\mathbf{y}^k)$ for $t > t_k$. Between observations $p(\mathbf{x}_t|\mathbf{y}^k)$, $t \in [t_k, t_{k+1}[$ satisfies Kolmogorov's forward equation (4.6), hence we get

$$dE^k(\varphi(\mathbf{x}_t)) = E^k(\frac{\partial\varphi}{\partial\mathbf{x}}\mathbf{f}(\mathbf{x}_t))\,dt + \frac{1}{2}\mathrm{tr}E^k(\mathbf{G}\mathbf{G}^T\frac{\partial^2\varphi}{\partial\mathbf{x}^2})\,dt \tag{4.15}$$

for $t \in [t_k, t_{k+1}[$, see (Jazwinski, 1970). Then the change in $E^k(\varphi(\mathbf{x}_t))$ is computed, when a new observation becomes available at time t_k using the difference equation (4.11) for the conditional density. Multiplying (4.11) by $\varphi(\mathbf{x})$ and integrating over $\mathbf{x}$, we have

$$E^k(\varphi(\mathbf{x}_k)) = \frac{E^{k-1}(\varphi(\mathbf{x}_k)p(\mathbf{y}_k|\mathbf{x}_k))}{E^{k-1}(p(\mathbf{y}_k|\mathbf{x}_k))} \tag{4.16}$$

Now the tools to determine the evolution of all the moments of the conditional density $p(\mathbf{x}_t|\mathbf{y}^k)$, $t \in [t_k, t_{k+1}[$ for all $k > 0$ are available. Consider the propagation of the conditional mean and covariance

$$\hat{\mathbf{x}}_{t|k} = E^k(\mathbf{x}_t) \tag{4.17}$$

$$\hat{\mathbf{P}}_{t|k} = E^k((\mathbf{x}_t - \hat{\mathbf{x}}_{t|k})(\mathbf{x}_t - \hat{\mathbf{x}}_{t|k})^T) = E^k(\mathbf{x}_t\mathbf{x}_t^T) - \hat{\mathbf{x}}_{t|k}\hat{\mathbf{x}}_{t|k}^T \tag{4.18}$$

The results can be obtained by setting $\varphi(\mathbf{x}) = \mathbf{x}$ and $\varphi(\mathbf{x}) = \mathbf{x}\mathbf{x}^T$ respectively in (4.15) and (4.16).

Theorem 4.1 *Assume the conditions required for the derivation of the conditional density in (4.6) and (4.11). Between observations, the* *conditional mean and covariance satisfy*

$$d\hat{\mathbf{x}}_{t|k}/dt = \widehat{\mathbf{f}(\mathbf{x}_t, t)} \tag{4.19}$$

$$d\mathbf{P}_{t|k}/dt = \widehat{\mathbf{x}_t\mathbf{f}^T} - \hat{\mathbf{x}}_{t|k}\widehat{\mathbf{f}^T} + \widehat{\mathbf{f}\,\mathbf{x}_t^T} - \widehat{\mathbf{f}}\,\hat{\mathbf{x}}_{t|k}^T + \widehat{\mathbf{G}\mathbf{G}^T} \tag{4.20}$$

for $t \in [t_k, t_{k+1}[$, *where* $\widehat{(\,\cdot\,)} \triangleq E^k(\,\cdot\,)$. *When a new observation arrives at* t_k *we have*

$$\hat{\mathbf{x}}_{k|k} = \frac{E^{k-1}(\mathbf{x}_k p(\mathbf{y}_k|\mathbf{x}_k))}{E^{k-1}(p(\mathbf{y}_k|\mathbf{x}_k))} \tag{4.21}$$

$$\mathbf{P}_{k|k} = \frac{E^{k-1}(\mathbf{x}_k\mathbf{x}_k^T p(\mathbf{y}_k|\mathbf{x}_k))}{E^{k-1}(p(\mathbf{y}_k|\mathbf{x}_k))} - \hat{\mathbf{x}}_{k|k}\hat{\mathbf{x}}_{k|k}^T \tag{4.22}$$

Predictions $\hat{\mathbf{x}}_{t|k}$ *and* $\mathbf{P}_{t|k}$, *with* $t > t_k$, *based on* $\mathbf{y}^k$, *also satisfy (4.19) and (4.20).*

PROOF: *is found e.g. in (Maybeck, 1982).*□

Notice that the equations in theorem 4.1 are not ordinary differential and difference equations. The right-hand sides of the equations involve expectations which require the whole conditional density for their evaluation. Apparently, in order to obtain a computationally realizable and practical predictor in the general nonlinear case, some approximations must be made. There are however two interesting special cases, where exact filtering is possible.

4.1.3 Deterministic Model

The first class is trivial, but so common that it is worth mentioning. It involves a *deterministic* general nonlinear state space model and an associated measurement equation with additive, uncorrelated Gaussian noise.

This corresponds to the Itô equation (4.1) with the term $\mathbf{G} \equiv 0$,

$$d\mathbf{x}_t = \mathbf{f}(\mathbf{x}_t, \mathbf{u}_t, \theta, t)dt. \tag{4.23}$$

The observations $\mathbf{y}_k$ are taken at discrete time instants, t_k

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k, \mathbf{u}_k, \theta, t_k) + \mathbf{e}_k \tag{4.24}$$

where $\mathbf{e}$ is a Gaussian white noise process $\mathbf{e}_k \sim N(0, \mathbf{S}_k(\theta))$. This class of models do not have a noise model and in this sense they may be called *white* box because the process is assumed to be deterministic. The parameters, though, are still unknown. In this case a stochastic framework is not needed at all. The filtering becomes trivial since $\hat{\mathbf{y}}_k = \mathbf{h}(\mathbf{x}_k, \mathbf{u}_k, \theta, t_k)$ with $\mathbf{x}_k$ given by the deterministic solution of (4.23).

The model (4.23) and (4.24) may be viewed as a degenerate *innovations representation*, because the innovations are given by the sequence $\{\mathbf{e}_k\}$ and the model is formulated explicitly as an additive function of them. Generally though models must be transformed into this form. Finding such a transformation may be difficult, and in the general case, amounts to solving the general filtering problem.

4.1.4 Linear Model

A second class of models allowing exactly filtering, is the common class of linear models. For this class of models the Kalman filter provides the exact solution for the filtering problem. Consider the model

$$d\mathbf{x}_t = \mathbf{A}(\mathbf{u}_t, \theta, t)\,\mathbf{x}_t\,dt + \mathbf{B}(\mathbf{u}_t, \theta, t)\,\mathbf{u}_t\,dt + \mathbf{G}(\theta, t)\,d\beta_t \tag{4.25}$$

$$\mathbf{y}_k = \mathbf{C}(\mathbf{u}_k, \theta, t_k)\,\mathbf{x}_k + \mathbf{D}(\mathbf{u}_k, \theta, t_k)\,\mathbf{u}_k + \mathbf{e}_k \tag{4.26}$$

There are different approaches leading to the model (4.25) and (4.26). The model may be formulated directly in this form, notice that we do not require linearity in the parameters. The model may typically be formulated as a linear model, but with coefficients varying according to some known external signal. Another approach leading to this class of models, is a linearization of the general Itô differential equation (4.1) and (4.2) around some reference signal $\mathbf{x}^*$. In this case the matrices are calculated by $\mathbf{A}(\mathbf{u}_t, \theta, t) = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\big|_{x=x^*}$, $\mathbf{B}(\mathbf{u}_t, \theta, t) = \frac{\partial \mathbf{f}}{\partial \mathbf{u}}\big|_{x=x^*}$ etc.

For the class of linear models (4.25) and (4.26) the Kalman filter provides the exact solution. The following equations are obtained for updating the state estimate:

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \epsilon_k \tag{4.27}$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k \mathbf{R}_{k|k-1} \mathbf{K}_k^T \tag{4.28}$$

$$\mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{C}^T \mathbf{R}_{k|k-1}^{-1} \tag{4.29}$$

The formulas for prediction of mean and covariance of the state-vector and observations are given by,

$$d\hat{\mathbf{x}}_{t|k}/dt = \mathbf{A}\,\hat{\mathbf{x}}_{t|k} + \mathbf{B}\,\mathbf{u}_t\,, \qquad t \in [t_k, t_{k+1}[\tag{4.30}$$

$$d\mathbf{P}_{t|k}/dt = \mathbf{A}\,\mathbf{P}_{t|k} + \mathbf{P}_{t|k}\mathbf{A}^T + \mathbf{G}\mathbf{G}^T\,, \qquad t \in [t_k, t_{k+1}[\tag{4.31}$$

$$\hat{\mathbf{y}}_{k+1|k} = \mathbf{C}\,\hat{\mathbf{x}}_{k+1|k} + \mathbf{D}\,\mathbf{u}_{k+1} \tag{4.32}$$

$$\mathbf{R}_{k+1|k} = \mathbf{C}\,\mathbf{P}_{k+1|k}\mathbf{C}^T + \mathbf{S} \tag{4.33}$$

The initial conditions are $\hat{\mathbf{x}}_{1|0} = \mu_0$ and $\mathbf{P}_{1|0} = \mathbf{V}_0$. The dependencies of time and external input of the matrices in the Kalman filter equations have been suppressed for convenience. This implementation of the Kalman filter thus involves the solution of a set of ordinary differential equations between each sampling instant. If, on the other hand the matrices $\mathbf{A}$, $\mathbf{B}$

and $\mathbf{G}$ are time invariant, then it is possible to find an explicit solution for (4.30) and (4.31), by integrating the equations over the time interval $[t_k, t_{k+1}[$ and assuming that $u_t = u_k$ in this interval, thus obtaining

$$\hat{\mathbf{x}}_{k+1|k} = \mathbf{\Phi}\,\hat{\mathbf{x}}_{k|k} + \mathbf{\Gamma}\,\mathbf{u}_k \tag{4.34}$$

$$\mathbf{P}_{k+1|k} = \mathbf{\Phi}\,\mathbf{P}_{k|k}\mathbf{\Phi}^T + \mathbf{\Lambda} \tag{4.35}$$

where the matrices $\mathbf{\Phi}$, $\mathbf{\Gamma}$ and $\mathbf{\Lambda}$ are calculated as,

$$\begin{aligned} &\mathbf{\Phi}(\tau) = e^{\mathbf{A}\tau}\ ; \qquad \mathbf{\Gamma}(\tau) = \int_0^\tau e^{\mathbf{A}s}\mathbf{B}\,ds\ ; \\ &\mathbf{\Lambda}(\tau) = \int_0^\tau \mathbf{\Phi}(s)\mathbf{G}\mathbf{G}^T\mathbf{\Phi}(s)^T\,ds \end{aligned} \tag{4.36}$$

and τ is the sampling time. This implementation of the Kalman filter thus involves the calculation of the exponential of a matrix. This calculation may be done once for a given set of parameters if the matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{G}$ are time invariant.

If the time dependence is slow compared to the dominating eigenvalues of the system, this implementation of the Kalman filter may also be used for time varying systems, by evaluating (4.36) for each sampling instant, assuming that $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{G}$ are constant within a sampling time. This solution requires less computations and is more robust than integrating (4.30) and (4.31), see (Moler & van Loan, 1978; van Loan, 1978). This point is also discussed in Graebe (1990b, p. 81).

External Inputs

In the derivation of (4.34) and (4.35) it is assumed that the external input is constant throughout the sampling interval, i.e. for $t \in [t_k, t_{k+1}[$ we have $\mathbf{u}_t \equiv \mathbf{u}_k$. This may be true in some cases, e.g. when the input is controlled. There is however in general a problem with external inputs. In the predictor the external input is assumed to be known, and in the continuous filter, this goes for all $t \in [t_0, T]$. Usually we have only discrete observations, at times t_k, of both inputs and outputs. One approach is to consider it as an continuous-discrete smoothing problem, that is to estimate $\mathbf{u}_t$, given $\mathbf{u}^k$ for $t < t_k$ We then have to impose certain assumptions, basically on the model for generating the input. A simple approach is to use linear interpolation for $\mathbf{u}_t$ through the sampling interval, hence

$$\mathbf{u}_t = \frac{t - t_k}{\tau}(\mathbf{u}_{k+1} - \mathbf{u}_k) + \mathbf{u}_k \; ; \qquad t \in [t_k, t_{k+1}[\tag{4.37}$$

where τ is the sampling time. When inserting this equation for $\mathbf{u}_t$, we obtain the following equation to replace (4.34),

$$\hat{\mathbf{x}}_{k+1|k} = \mathbf{\Phi}\, \hat{\mathbf{x}}_{k|k} + \mathbf{\Gamma}\, \mathbf{u}_k + \mathbf{\Upsilon}\, (\mathbf{u}_{k+1} - \mathbf{u}_k) \tag{4.38}$$

where $\mathbf{\Phi}$ and $\mathbf{\Gamma}$ are the same as before, and

$$\mathbf{\Upsilon}(\tau) = \int_0^{\tau} e^{\mathbf{A}s} \mathbf{B} \tfrac{\tau - s}{\tau} \, ds \tag{4.39}$$

Notice that this approach is very simple to implement as an integrated part of the discretization procedure, see Chapter A for details about implementations. Also for most cases the assumption of linear interpolation between consecutive observations of inputs is more realistic than that of a constant value between the samples.

4.1.5 Other Criteria of Optimality

In the previous section the filter was derived in a Bayesian manner by generating explicit recursions for the conditional probability density for the states, conditioned on the entire measurement history. In the linear model case we get the Kalman filter, and the conditional density is Gaussian. The optimal state predictions were given as the conditional mean.

Theorem 4.2 *The estimate that minimizes* $E((x_t - \hat{x}_t)^T W (x_t - \hat{x}_t))$, *where* $W \geq 0$ *is some weight matrix, is called the minimum variance estimate. Let the estimate,* $\hat{x}_t$ *be a functional on* y^k *for* $t \geq t_k$. *Then the minimum variance estimate is the conditional mean.*

Proof: *see (Jazwinski, 1970).* □

There are other criteria of optimality that might be interesting for an estimation problem, e.g. the maximum (mode) or the median of the a posteriori conditional density. When the probability density is Gaussian, the mean is also the mode (= MAP estimate) and median of the density function, this holds for all symmetric and unimodal (only one peak) densities.

It is possible to derive formulas for the evolution of the mode and the covariance similar to those in section 4.1.2. During time propagations from t_k to t_{k+1} the mode is defined by

$$dp(\xi, t|y^t)/d\xi\big|_{\xi = \hat{x}_{map}} \equiv 0 \tag{4.40}$$

provided the second derivative is positive definite to assure that the mode is well defined, it is also required that the mode is unique. The time derivative in (4.40) must also be identically zero. Also for the mode estimator we obtain an infinite dimensional estimator, and as before approximations are required to generate an implementable finite dimensional filter. It seems, from the literature, that the practical experiences with this alternative formulation of the predictor is little, (Jazwinski, 1970; Maybeck, 1982).

4.2 Approximate Filters

We will now concentrate on computing the conditional mean and covariance in the nonlinear filter. It was shown in previous sections that generally the calculations of these moments requires access to the whole conditional density function. In general the conditional density function can not be characterized by a finite set of parameters (e.g. its moments). An important exception is of course the linear filtering problem, in which case the conditional density is Gaussian, and therefore completely determined by its mean and covariance. We may consider the mean and covariance as the *hyper-state* of the linear filter. In the nonlinear case, the filter state is essentially infinite dimensional.

It was shown in section 4.1.1 that the calculation of the conditional density involve solution of partial differential equations, which is infeasible for computation. Thus we are forced to consider approximations to the conditional density function. The approach is to *parameterize* the conditional density via a finite and small set of parameters. The nonlinear filter would then consists of equations of evolution for these parameters, which would comprise the state of the filter. If we could find a finite set of parameters that completely determine the conditional density, these parameters would be sufficient statistics. Unfortunately, it is virtually impossible to find sufficient statistics for nonlinear problems, the linear case being a unique exception.

There are a number of possible means of approximately parameterizing the conditional density function. A valid approach is to express the density in terms of a complete orthogonal series and then truncating the series at a specified order, (Maybeck, 1982). However, here will be focused on the parameterizing via moments.

4.2.1 First Order Filters

The simplest form of approximations in nonlinear filtering is based on the Kalman filter applied to some linearization of the model. Filters of this class are called first order approximations. We presented the Kalman filter applied to a linear model in section 4.1.4. Here we consider the linear model as generated from a linearization of the Itô equation using (4.44). In this case the matrices are calculated by $\mathbf{A}(\mathbf{u}_t, \theta, t) = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\big|_{x=x^*}$, $\mathbf{B}(\mathbf{u}_t, \theta, t) = \frac{\partial \mathbf{f}}{\partial \mathbf{u}}\big|_{x=x^*}$ etc., where $\mathbf{x}^*$ is some reference signal. This reference trajectory is chosen as the one obtained by setting the noise to zero in the nonlinear model

$$d\mathbf{x}_t^* = \mathbf{f}(\mathbf{x}_t^*, \mathbf{u}_t, \theta, t)dt \tag{4.41}$$

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k^*, \mathbf{u}_k, \theta, t_k) \tag{4.42}$$

and integrating these deterministic equations over each sample period. Linearizing about $\mathbf{x}^*$ yields the linear perturbations model, in the form of a linear time varying state-space model. This method, called the *linearized Kalman filter*, will only converge if the noise levels are sufficiently small. This is due to the assumption of zero noise in the nonlinear model for calculating the reference trajectory.

A better choice for the linearization trajectory, is to use the current estimate of the state. Linearizing about it at every sampling time and applying a Kalman filter to the resulting linearized model yields the algorithm known as *extended Kalman filter*. In this manner, one enhances the validity of the assumption that deviations from the "true" trajectory are small enough to allow linear perturbation techniques to be employed with adequate results.

Performance improvement for the extended Kalman filter may be obtained by local iterations (over a single sample period) on nominal trajectory redefinition and subsequent relinearization. If we iterate on the equations for the measurement update, by replacing equation (4.27) with the following iterator

$$\eta_i = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{y}_k - \mathbf{h}(\eta_{i-1}, t_k) - \mathbf{C}\,(\hat{\mathbf{x}}_{k|k-1} - \eta_{i-1})) \tag{4.43}$$

with $\mathbf{C} = (\partial\mathbf{h}/\partial\mathbf{x})\big|_{x=\eta_{i-1}}$, and $\mathbf{K}_k = \mathbf{K}_k(\eta_{i-1})$, iterated for $i = 1, \cdots, l$, starting with $\eta_0 = \hat{\mathbf{x}}_{k|k-1}$, and terminating with the result $\hat{\mathbf{x}}_{k|k} = \eta_l$, we have the algorithm called *iterated extended Kalman filter*. The filter just described addresses the problem of nonlinearities by reevaluating $\mathbf{h}$ and $\mathbf{C}$ to achieve a better $\hat{\mathbf{x}}_{k|k}$. This will also improve estimation over future intervals because of improved succeeding reference trajectories. It is also possible to improve the reference trajectory backward in time once the measurement y_k is taken, by applying smoothing techniques backward to time t_{k-1}. Incorporating such a local iteration into the extended Kalman filter structure yields what is termed the *iterated linearized filter-smoother*. It has been demonstrated by simulations, that these filters are effective in nonlinear problems, but both of these iterated filters produce biased estimates in general, see e.g. (Ljung, 1979). However, as the error variance becomes small, so does the bias in the estimate, see (Jazwinski, 1970, pp. 278–281).

Another filter to be considered in this section is based on statistical linearization, and therefore called the *statistically linearized filter*. This filter can be applied if the diffusion term is independent of the state, $\mathbf{x}_t$, i.e. the term

$$\mathbf{G} = \mathbf{G}(\theta, t) \tag{4.44}$$

in (4.1) is only a function of θ and t. Contrary the other filters mentioned in this section this filter does not imply series representations of $\mathbf{f}$ and $\mathbf{h}$. For this reason, it is a method that does not require $\mathbf{f}$ and $\mathbf{h}$ to be differentiable, thereby admitting such important nonlinearities as saturation. This advantage is gained at the expense of requiring evaluation of conditional expectations, i.e. knowing the entire conditional density, according to (4.17). Typically, the density is approximated as Gaussian, and the resulting implementable algorithm often has better characteristics than those based on truncated series expansions of $\mathbf{f}$ and $\mathbf{h}$ about the conditional mean approximate estimate, see (Maybeck, 1982, pp. 243–245). Consider approximating $\mathbf{f}(\mathbf{x}_t, t)$ by a linear approximation of the form

$$\mathbf{f}(\mathbf{x}_t, t) = \mathbf{f}_0(t) + \mathcal{F}(t)\,\mathbf{x}_t + \epsilon_t \tag{4.45}$$

which has the minimum mean square error

$$J = E(\epsilon_t^T \boldsymbol{W} \epsilon_t | \boldsymbol{z}^{k-1}) \tag{4.46}$$

for all $t \in [t_{k-1}, t_k[$, where $\boldsymbol{W} \geq \mathbf{0}$ is a weighting matrix. Calculating the partial derivatives of (4.46), with respect to $\mathbf{f}_0(t)$ and $\mathcal{F}(t)$ and setting them to zero yields, using the notation of Theorem 4.1

$$\mathbf{f}_0(t) = \widehat{\mathbf{f}} - \mathcal{F}(t)\,\hat{\mathbf{x}}_t \tag{4.47}$$

$$\mathcal{F}(t) = (\widehat{\mathbf{f}\mathbf{x}_t^T} - \widehat{\mathbf{f}}\hat{\mathbf{x}}_t^T)\,\mathbf{P}_{t|t}^{-1} \tag{4.48}$$

with $\mathbf{P}_{t|t}$ being the conditional covariance of $\mathbf{x}_t$. $\mathcal{F}(t)$ is close related to *describing function theory* for approximating nonlinearities. In the scalar case (4.48), with zero mean, becomes $\widehat{fx}/\widehat{x^2}$, which is the describing function gain for an odd-function nonlinearity (such as a symmetric saturation), see (Gelb, 1974, pp. 204–220). Using similar approximations for the measurement equation yields

$$\mathbf{h}(\mathbf{x}_k, t_k) \cong \mathbf{h}_0(t_k) + \mathcal{H}(t_k)\,\mathbf{x}_k \tag{4.49}$$

with the coefficients statistically optimized to get

$$\mathbf{h}_0(t_k) = \widehat{\mathbf{h}} - \mathcal{H}(t_k)\,\hat{\mathbf{x}}_{k-1} \tag{4.50}$$

$$\mathcal{H}(t_k) = (\widehat{\mathbf{h}\mathbf{x}_k^T} - \widehat{\mathbf{h}}\hat{\mathbf{x}}_{k-1}^T)\,\mathbf{P}_{k|k-1}^{-1} \tag{4.51}$$

The issue now is the computing of $\widehat{\mathbf{f}}$, $\widehat{\mathbf{h}}$, $\mathcal{F}(t)$ and $\mathcal{H}(t_k)$. They all depends upon the conditional probability density function of $\mathbf{x}$, which is generally not available. We therefore assume the density is Gaussian. Since this density is completely defined by its mean and covariance, both of which are part of the computation in the filtering algorithm, it will be possible to compute all the conditional expectations introduced in (4.48) and (4.51). We obtain the statistically linearized filter, with the following equations for the time propagation

$$d\hat{\mathbf{x}}_{t|k}/dt = \widehat{\mathbf{f}}(\mathbf{x}_{t|k}, t)\;, \qquad t \in [t_k, t_{k+1}[\tag{4.52}$$

$$d\mathbf{P}_{t|k}/dt = \mathcal{F}(t)\mathbf{P}_{t|k} + \mathbf{P}_{t|k}\mathcal{F}^T(t) + \mathbf{G}\mathbf{G}^T\;, \qquad t \in [t_k, t_{k+1}[\tag{4.53}$$

with $\mathcal{F}(t)$ given by (4.48), and the conditional expectations involved calculated assuming $\mathbf{x}_t$ to be Gaussian with mean $\hat{\mathbf{x}}_{t|k}$ and covariance $\mathbf{P}_{t|k}$. The measurement update at time t_k is given by

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathcal{H}^T(t_k)[\mathcal{H}(t_k)\mathbf{P}_{k|k-1}\mathcal{H}^T(t_k) + \mathbf{S}]^{-1} \tag{4.54}$$

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{y}_k - \widehat{\mathbf{h}}) \tag{4.55}$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k\mathcal{H}(t_k)\mathbf{P}_{k|k-1} \tag{4.56}$$

with $\mathcal{H}(t_k)$ given by (4.51), and the conditional expectations calculated as though $\mathbf{x}_k$ were Gaussian with mean $\hat{\mathbf{x}}_{k|k-1}$ and covariance $\mathbf{P}_{k|k-1}$. Structurally, the equations for the gain and covariance are the same as those for the extended Kalman filter, but with $\mathcal{F}(t)$ replacing $(\partial \mathbf{f}/\partial \mathbf{x})\big|_{x=\hat{x}_{t|k}}$ and $\mathcal{H}(t_k)$ replacing $(\partial \mathbf{h}/\partial \mathbf{x})\big|_{x=\hat{x}_{k|k-1}}$. The computational requirements of the statistically linearized filter may be greater than for filters derived from Taylor series expansions of the nonlinearities because the expectations must be performed over the assumed Gaussian density of $\mathbf{x}$. However, Monte Carlo simulations have demonstrated that the performance advantages offered by statistical linearization may make the additional computations worthwhile, see (Gelb, 1974).

4.2.2 Second Order Filters

There are mainly two different approaches, for approximating the exact solution given in section 4.1.2, leading to filter expressions belonging to the class of second order filters. One might assume that the conditional density is nearly symmetric so that third and higher order odd central moments are essentially zero, and also that it is concentrated sufficiently closely about the mean that the fourth and higher order even central moments are small enough to be neglected, this leads to the *truncated second order filter*. The other approach is to assume that the conditional density is nearly Gaussian, so that third and higher order odd central moments are again essentially zero, and the fourth and higher order even central moments can be expressed in terms of the covariance, this leads to the *Gaussian second order filter*. Typically, sixth and higher order even moments are also assumed small enough to be neglected. For both filters further approximations of the functions $\mathbf{f}$ and $\mathbf{h}$ are applied, via a truncated Taylor series expanded about the current state estimate. However there are also *assumed density filters* that do not require such series approximations, corresponding to the statistically linearized filter in the previous section.

It should be noted, that the Gaussian approximation in general has a wider range of validity than the truncated approximation. The difference between the two approaches is the assumption about the fourth moment, in the truncated second order filter it is neglected, while in the Gaussian case it is approximated as

$$E^{\tau}((x_i - \hat{x}_i)(x_j - \hat{x}_j)(x_k - \hat{x}_k)(x_l - \hat{x}_l)) = P_{jk}P_{il} + P_{jl}P_{ik} + P_{kl}P_{ij} \qquad (4.57)$$

where $\mathbf{P}$ is the conditional covariance matrix. For both second order filters, we get the equations for the time propagation, between observations

$$d\hat{\mathbf{x}}_{t|k}/dt = \mathbf{f}(\hat{\mathbf{x}}_{t|k}, t) + \frac{1}{2}\mathbf{P}_{t|k}\frac{\partial^2 \mathbf{f}(\hat{\mathbf{x}}_{t|k}, t)}{\partial \mathbf{x}^2}, \qquad t \in [t_k, t_{k+1}[\qquad (4.58)$$

$$d\mathbf{P}_{t|k}/dt = F\mathbf{P}_{t|k} + \mathbf{P}_{t|k}F^T + \widehat{\mathbf{G}\mathbf{G}^T}, \qquad t \in [t_k, t_{k+1}[\qquad (4.59)$$

where $F = \partial \mathbf{f}(\hat{\mathbf{x}}_{t|k}, t)/\partial \mathbf{x}$. The evaluation of $\widehat{\mathbf{G}\mathbf{G}^T}$ differ for the two filters, due to the relation (4.57). The predictions of the output is needed for the residual generation

$$\hat{\mathbf{y}}_{k|k-1} = \mathbf{h}(\hat{\mathbf{x}}_{k|k-1}, t_k) + \frac{1}{2}\mathbf{P}_{k|k-1}\frac{\partial^2 \mathbf{h}(\hat{\mathbf{x}}_{k|k-1}, t_k)}{\partial \mathbf{x}^2} \qquad (4.60)$$

For measurement updating at sample time t_k, one could attempt to approximate the expectation integrations inherent in (4.21) and (4.22). It turns out, that a better approximation is to assume that the conditional mean and covariance can be expressed as a power series in the innovations, (Jazwinski, 1970; Maybeck, 1982). For computational tractability, this power series is truncated at first order terms

$$\hat{\mathbf{x}}_{k|k} = \mathbf{a}_0 + \mathbf{a}_1(\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}) \qquad (4.61)$$

$$\mathbf{P}_{k|k} = \mathbf{b}_0 + \sum_{i=0}^{s} \mathbf{b}_{1,i}\{\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}\}_i \qquad (4.62)$$

where $\mathbf{a}_0$ in an n-vector, $\mathbf{a}_1$ is $n \times s$, $\mathbf{b}_0$ is $n \times n$ and the matrices $\mathbf{b}_{1,i}$ are $n \times n$. The calculation of the terms $\mathbf{a}_0$, $\mathbf{a}_1$, $\mathbf{b}_0$, $\mathbf{b}_{1,i}$ for the two filters may be found in e.g. (Maybeck, 1982).

Setting $\forall i : \mathbf{b}_{1,i} \equiv \mathbf{0}$ in equation (4.62) of the full second order filters, yields the *modified Gaussian second order filter* and *modified truncated second order filter* respectively. Hereby the computations are reduced significantly compared to the full second order filter, but still maintaining correction terms compared to first order filters. When $\mathbf{G}$ is not a function of $\mathbf{x}$ and the second partial derivatives $\partial^2\mathbf{f}/\partial\mathbf{x}^2$ and $\partial^2\mathbf{h}/\partial\mathbf{x}^2$ are neglected in the second order filters, they reduce to the extended Kalman filter. Similarly it is observed that all the approximate nonlinear filters reduce to the Kalman filter when the dynamics and observations are linear. If we use equation (4.58) and (4.60) from the second order filter, but maintain the covariance expression from the first order filter, we obtain a *first order filter with bias correction*. The last terms of these two equations are called the bias correction terms. With this compromise, we may benefit essential from the second order filtering, but without the significant additional computational burden of the full second order filter.

4.2.3 Convergence of the approximate filters

The important questions for all the mentioned approximate filters are about their convergence properties and bias of estimates. It was shown by Ljung (1979) that for the EKF when used for parameter estimation in linear models convergence is not guaranteed. By adding a bias correction term Ljung showed that this new filter is convergent for a general class of models. The EKF was designed for the case $\mathbf{G} = \mathbf{G}(\theta, t)$, i.e. the diffusion term is not dependent of the state vector. When the state vector is extended with the parameter vector the diffusion term becomes dependent of the augmented state vector and in general the estimates are biased when the EKF without bias correction is used.

In general it is not possible to guarantee convergence of any of the approximate filters. The convergence properties and bias of the estimates will depend upon the type and magnitude of the nonlinearities. For the methods implying series representation of f, G and h it is required that they are differentiable, whereas this is not a requirement for the assumed density filters. The benefit from including higher order terms increases with the magnitude of the nonlinearity. But the significance of including extra terms is lowered if the level of the noise is increased, i.e. biases due to neglecting higher order effects are masked the substantial spreading of the density functions, see (Maybeck, 1982).

4.3 Summary

In this chapter we have described the nonlinear filtering problem for a continuous time system, with discrete time measurements. Essentially we are interested in the first two conditional moments of the output. They are to be used in the calculation of the Gaussian likelihood function for the parameter estimation, described in chapter 3.

The formulas for the exact evaluation of these moments has been derived. This is mainly a conceptual solution of the problem, since only in special (but important) cases the exact solution is computationally feasible. In the general case one has to go for an approximate solution to the filtering problem. In the chapter a number of approximate filters are mentioned ranging from simple first order filters to full second order filters. From a practical and computational point of view it is probably some of the "compromise" filters that are most interesting, e.g. the *extended Kalman filter with bias correction* and *modified Gaussian second order filter*. It should also be mentioned that the filters with an assumed density (usually Gaussian) of the state vector, have a wider range of validity in practical applications.

In general there is always a tradeoff between the validity of a certain approximative filter and the computational burden involved. The approximate filters have been found useful in a number of applications (Jazwinski, 1970; Maybeck, 1982), but in general it is not possible to guarantee the convergence for any of the approximate filters. The stability and convergence properties for the filter in a particular application must be tested by Monte Carlo simulations. It will also be dependent of the character of the nonlinearities in a particular application which of the approximate filters is the most favorable to use for that application.

Chapter 5

Model Validation

A number of methods involved for validation of physical models or grey box models are closely related to the techniques used for black box models. However, the advantage of physical models is the prior information about the model structure and the parameters. Compared to a traditional black box model, the physical model often has a nonlinear model structure assumed to be partly known a priori. This also imply that the class of alternative model structures is often limited. The a priori information about the parameters of the physical model can either be in terms of an allowed range of the parameters or even a prior distribution of the parameters. The prior distribution can either express a prior belief or be determined from data from a previous experiment. The introduction of prior information implies a special demand on the statistical tests, which have to be evaluated in the domain of the claimed information. For some of the methods this can be done in a formal way using a Bayesian approach where prior information is specified as a prior distribution of e.g. the parameters. A less formal but still important validation for physical models is to confront the estimated parameters and their covariance with the prior knowledge.

The intended use of the model is important in the validation. If the purpose of the model is reflected by the experimental conditions $\mathcal{X}$, under which the data is collected, then a given model should be valid under these experimental conditions. So the model validation is dependent upon the experimental conditions, see Chapter 6. By taking this approach, we omit the difficulty of defining a specific class of "purposive" models for the model validation as used by Bohlin (1989). If, however, a class of purposive models can be defined, it may be adequate to evaluate the resulting models in a Bayesian framework and define the purposivity by loss functions. The separation between purposive and non-purposive models can then be achieved via an appropriate choice of loss function. Hence, the validation procedure can also be studied using a Bayesian approach.

In the following different tests concerning model validation are discussed including both time and frequency domain methods. Classical methods based on repeated sampling are compared with Bayesian techniques. Model validation of grey box models are discussed by e.g. (Holst, Holst, Madsen, & Melgaard, 1992; Bohlin, 1978; Graebe, 1990a).

5.1 Test for Model Structure

The tests typically compare alternative model structures. This type of tests are usually applied for checking over-fitting and possible model reduction.

5.1.1 Likelihood based tests

For the kind of physical models that are considered here, the structural information is formulated in continuous time. However, the data are given in discrete time. For the actual identification of the embedded parameters in the continuous time model, the family of likelihood based methods are adequate because they are invariant over complicated transformations like the transformation from continuous time models to discrete time models, see Chapter 3.

Consider the problem of testing the hypothesis:

$$H_0 : \theta \in M_0 \quad \text{against} \quad H_1 : \theta \in M_1 \tag{5.1}$$

where $M_0 \subset M_1$ and M_0 and M_1 represent two model structures. A test based on the *likelihood ratio*

$$\lambda = \mathcal{L}(\hat{\theta}_0)/\mathcal{L}(\hat{\theta}_1) \tag{5.2}$$

where $\mathcal{L}(\hat{\theta}_0)$ is the likelihood function for the maximum likelihood estimate $\hat{\theta}_0$ under H_0 and $\mathcal{L}(\hat{\theta}_1)$ is the likelihood function for the maximum likelihood estimate $\hat{\theta}_1$ under H_1 can be a "high powered" test for the hypothesis testing, see (Kendall & Stuart, 1979; Goodwin & Payne, 1977). In a few cases it is possible to derive exact distributions of the test quantity under H_0. When this is not possible the following large sample test can be used. Under weak assumptions the following holds asymptotically under H_0

$$LR = -2\log\lambda \sim \chi^2(r) \tag{5.3}$$

where λ is the likelihood ratio and the degrees of freedom is $r = \dim(M_1) - \dim(M_0)$.

It is known from Section 3.3.2 that the maximum likelihood estimator, $\hat{\theta}$ is asymptotically normally distributed with mean θ and covariance $\mathbf{M}^{-1}$, where $\mathbf{M}$ is the Fisher information matrix. This can be used for asymptotic tests on individual parameters. From the covariance estimate it is also possible to calculate the correlation matrix of the parameters. Inspection of this matrix may indicate if parameters are highly correlated, such that one of them can be removed from the model.

Consider the testing of a general hypothesis of the form

$$H_0 : \mathbf{R}(\theta) = 0 \quad \text{against} \quad H_1 : \mathbf{R}(\theta) \neq 0$$

where the restriction $\mathbf{R}$ is a k-dimensional vector function. In order to derive the Wald test we have the following result analogous to that of the ML estimator

$$(\mathbf{R}(\hat{\theta}) - \mathbf{R}(\theta)) \xrightarrow{d} N(0, F^T \mathbf{M}^{-1} F) \tag{5.4}$$

where $F = \partial \mathbf{R}^T / \partial \theta$, and $\hat{\theta}$ is the unrestricted maximum likelihood estimator, i.e. the estimator under H_1. The *Wald test* statistic is given by

$$W = \mathbf{R}(\hat{\theta})^T [F^T \mathbf{M}^{-1} F]^{-1} \mathbf{R}(\hat{\theta}) \tag{5.5}$$

where F is evaluated at $\theta = \hat{\theta}$. Under H_0 the W-statistic is asymptotically χ^2-distributed with k degrees of freedom. The likelihood ratio test requires a maximum likelihood estimation for both the restricted and unrestricted estimators, while the Wald test is based only on the unrestricted estimator.

A third likelihood based test is the *Lagrange multiplier* test which is also called score test. The test statistic is given by

$$LM = \mathbf{S}(\hat{\theta}_0)^T \mathbf{M}(\hat{\theta}_0)^{-1} \mathbf{S}(\hat{\theta}_0) \tag{5.6}$$

where $\mathbf{S} = \partial \log \mathcal{L} / \partial \theta$, which is called the score statistics for θ, and $\hat{\theta}_0$ denotes the maximum likelihood estimate of θ under H_0. $\mathbf{S}$ as well as $\mathbf{M}$ are computed in the alternative model but still evaluated for the parameter values corresponding to the null hypothesis, i.e. the parameters in the alternative model do not have to be estimated. It can be shown that the test statistic LM is asymptotically chi-square distributed. The LM test may be used for choosing among two nested model structures, corresponding to the hypothesis (5.1). The information matrix and score statistics are calculated for the extended model M_1, but evaluated at the ML estimate $\hat{\theta}_0$ of the restricted model M_0. In this case LM is asymptotically chi-square distributed with $r = \dim(\theta_1) - \dim(\theta_0)$ degrees of freedom. H_0 is rejected if $\mathrm{LM} > \chi^2_\alpha(r)$ for some level α, where α specifies the maximal probability of rejecting H_0 when it is true. As opposed to the LR test this test does not require explicit computation of $\hat{\theta}_1$ hence, it does not require over-fitting (Bohlin, 1978). If for example a linear AR structure is to be tested against various nonlinear alternatives such as e.g. bilinear, threshold or exponential autoregressive models, the test statistic reflects the properties of the alternative, cf. (Tong, 1990). The test can also be used to evaluate the necessity of model extensions.

There are also other applications of the LM test, Graebe (1990b, pp. 93–101) uses the test in cross validation of the model using an independent data set (see Section 5.4). He also uses the LM test for validation of the approximations involved in the estimation procedure. This question is of specific interest for physical models, which are often of nonlinear nature, hence involving a complex/approximative estimation procedure. As discussed in Chapter 4 it is not practically feasible to use exact filtering for a general nonlinear model. Hence, some approximative filter is involved for producing the residuals for the ML (or MAP) estimation. The question is whether the error introduced by the approximations are sufficiently small.

The principle of the test is the following. First the parameters of the model are estimated from the measured data, then the same model with the estimated parameters are used for simulating new data (using the methods for stochastic simulation described in Section 2.5). Now the test statistic is calculated from the simulated data with the same model and parameters estimated from the real data. LM is tested in a chi-square distribution with $r = \dim(\theta)$ degrees of freedom. If the test rejects H_0 then the errors from the approximations involved in the estimation procedure are not sufficiently small on the chosen level of α. The test is of course based on the fact that there are only small approximations involved for simulating the model. This is a reasonable requirement according to the discussion in Section 2.5.

The LR test, the Wald test and the LM test all have an asymptotic chi-square distribution with the same degrees of freedom. However, according to Rao (1973) none of the tests is in general uniformly most powerful.

5.1.2 Information Criteria for Order Selection

If order selection is a part of the identification procedure, the likelihood based tests discussed above readily are applied to the problem. One approach is to measure the marginal improvement by a test quantity, as the LR test, Wald's test or the LM test. Another possibility is to consider various extensions of the maximum likelihood procedure by assigning a cost for high complexity.

Average Information Criterion

This criterion can be derived by minimizing the Kullback-Leibler mean information criterion (Akaike, 1976). The resulting criterion

$$\mathrm{AIC} = 2n - 2\log\mathcal{L}(\theta, \mathbf{y}^N) \tag{5.7}$$

is thus a slight but important modification of the log-likelihood that penalizes large $n = \dim(\theta)$. The optimal order n is the one that minimizes the AIC criterion.

Bayesian Information Criterion

It has been pointed out in Schwarz (1978) that the AIC criterion is not a consistent estimator of the model order, it has a tendency to overestimate n. Using Bayesian arguments another criterion for choosing the model which is a posteriori most probable can be derived. This results in the criterion

$$\mathrm{BIC} = n\log N - 2\log\mathcal{L}(\theta, \mathbf{y}^N) \tag{5.8}$$

where N is the number of observations. As opposed to the AIC criterion the BIC criterion gives a consistent estimate of the model order at least for some type of models (Lütkepohl, 1985).

Final Prediction Error

For scalar autoregressive models it can be shown that the estimated mean square prediction error of the process is

$$\mathrm{FPE} = \frac{N+n}{N-n}\,\hat{\sigma}^2 \tag{5.9}$$

where $\hat{\sigma}^2$ is the maximum likelihood estimate of the variance of the driving white noise sequence, see (Brockwell & Davis, 1987). The FPE criterion is asymptotically equivalent to the AIC criterion and thus determines the same optimal n for large N.

5.2 Residual Analysis

The purpose of the residual analysis is primarily to check whether any obtained information contradicts the assumptions upon which the models and methods are build. But secondly we are interested in any information that indicates in which direction to develop or extend the models. There are a number of statistical tests available for these purposes.

5.2.1 Tests for Independence

It is important always to plot the residuals, and judge the graph for trends and inhomogeneity of the variance. From the plot of standardized residuals it is possible to check the data for outliers and bad data, which is good practice even if some robust norm was used for the identification, see Section 3.5. The next step is to check the sample autocorrelation function of the residuals. Instead of checking the individual values of the autocorrelation function, it is possible to pool the information into a single statistic, and perform a *Portmanteau Test*, see e.g. (Ljung & Box, 1978). In addition to the tests based on the sample autocorrelation function there are a number of tests for checking the hypothesis of the residuals being an i.i.d. sequence, such as e.g. *Test Based on Turning Points*, *Difference-Sign Test* and *Rank Test*, see (Brockwell & Davis, 1987).

5.2.2 Kolmogorov-Smirnov test

The *Kolmogorov - Smirnov test* is designed for testing hypotheses concerning equality between e.g. an assumed, F_0, and an empirical distribution, F_n. The test statistic is

$$D_n = \sup_x |F_n(x) - F_0(x)|$$

which has a known distribution under F_0. F_n is the empirical distribution of $x_1, \cdots, x_n$. A test on level α is therefore given by the critical area $C = \{(x_1, \cdots, x_n) | D_n > c\}$, where c is evaluated from $\Pr\{D_n > c | F = F_0\} = \alpha$, see (Kendall & Stuart, 1979). The test has a number of advantages, compared to other methods. It is exact and easy to apply, since the probability for *no* value exciting the confidence limits is computed.

The information basis for the grey box modelling may be formulated as prior distributions of e.g. data or disturbances. The modelling results in a posteriori distributions for data as well as disturbances, which can be tested against the prior belief, cf. (Holst et al., 1992).

Tests for distribution may also be used as a means for checking the whiteness of the residuals as in the black box case. The cumulated periodogram for the residual sequence has the same properties as a distribution function, and is tested against the distribution of white noise, cf. (Brockwell & Davis, 1987). There is no need for (troublesome) smoothing in the frequency domain, when using this test.

5.2.3 Cross Spectra

If the residuals are white noise, then the model is in good agreement with the true system. Another interesting question is whether the residuals are independent of inputs. If not, then there is more information contained in the output that originates from the input than explained by the current model. Independence may be tested using the sample cross covariance function, see e.g. (Box & Jenkins, 1976).

The cross spectrum, given as the Fourier transform of the cross covariance function between residuals and inputs is primarily used as a diagnostic tool, to indicate how further improvement is possible.

A useful quantity derived from the cross spectrum, $\Gamma_{\epsilon u}(k)$, is the coherency spectrum

$$\kappa^2_{\epsilon u}(\omega) = \frac{|\Gamma_{\epsilon u}(\omega)|^2}{\Gamma_\epsilon(\omega)\Gamma_u(\omega)} \tag{5.10}$$

where Γ_ϵ and Γ_u are the auto covariance functions of $\{\epsilon_t\}$ and $\{u_t\}$ respectively. $\kappa^2_{\epsilon u}(\omega) \in [0, 1]$ can be interpreted as a non dimensional measure of the correlation between two time series at a certain frequency.

An alternative way of inspecting the correlation between residuals and multiple inputs is by the multiple coherency spectrum. Consider the sequence of residuals $\{\epsilon_t\}$ and q different inputs $\{u_{1,t}\}, \{u_{2,t}\}, \cdots, \{u_{q,t}\}$ with their means subtracted. It is then possible to separate the variation of $\{\epsilon_t\}$ the following way

$$\epsilon_t = \sum_k h_{\epsilon,1} u_{1,t-k} + \cdots + \sum_k h_{\epsilon,q} u_{q,t-k} + Z_t \tag{5.11}$$

In other words, a linear model for the residuals based on the inputs is fitted, and a new set of residuals $\{Z_t\}$ obtained. The spectrum of the noise process $\{Z_t\}$ is given by

$$\Gamma_z(\omega) = \Gamma_\epsilon(\omega) - \sum_{i=1}^{q} H_{\epsilon i} \Gamma_{\epsilon u_i}(\omega) \tag{5.12}$$

which can also be written as

$$\Gamma_z(\omega) = \Gamma_\epsilon(\omega)[1 - \kappa^2_{\epsilon 12 \cdots q}(\omega)] \tag{5.13}$$

where $\kappa^2_{\epsilon 12 \cdots q}$ is the squared multiple coherency spectrum of the output process and the q input processes. This quantity measures the proportion of the residual spectrum which can be predicted linearly from the inputs at the different frequencies. It is a useful diagnostic tool for model improvement and assessment of the partial prior knowledge.

5.2.4 Bispectra

The bispectrum is the frequency domain representation of the third order moment of the stochastic process $\{x_t\}$.

$$M_3(u, v) = E[(x_t - \mu)(x_{t+u} - \mu)(x_{t+v} - \mu)] \tag{5.14}$$

It is the natural extension to the usual second order representation of the spectrum, and is useful when dealing with non Gaussian/non linear processes. The bispectrum of the stationary process $\{x_t\}$ is defined as the two dimensional Fourier transform of the third moment

$$\Gamma_3(\omega_1, \omega_2) = \sum_{u,v} M_3(u, v) \exp(-i2(u\omega_1 + v\omega_2)) \tag{5.15}$$

where $-\pi \leq \omega_1, \omega_2 < \pi$. By inserting an estimate of (5.14) in (5.15) we obtain an estimate of the spectrum which is called the raw spectrum. This estimate is central but not consistent, and we have to smooth the estimate with a lag kernel, in order to get a consistent estimate.

For a linear and Gaussian process we have $\Gamma_3(\omega_1, \omega_2) = 0$ in the whole frequency plane. There is an approximative test, cf. (Subba Rao & Gabr, 1984), for the hypothesis

$$\begin{aligned} H_0 : &\quad \forall \omega_1, \omega_2 : \Gamma_3(\omega_1, \omega_2) = 0 \\ H_1 : &\quad \exists \omega_1, \omega_2 : \Gamma_3(\omega_1, \omega_2) \neq 0 \end{aligned}$$

If H_0 is accepted, then the stochastic process $\{x_t\}$ is considered linear and Gaussian. If on the other hand H_0 is rejected, the process can be either non Gaussian and/or nonlinear. A second step is proposed in Subba Rao & Gabr (1984) to test the hypothesis:

$$\begin{aligned} H_0 : &\quad x_t \text{ is linear and } M_3 = 0 \\ H_1 : &\quad \text{all alternatives} \end{aligned}$$

where M_3 is the third order moment given in (5.14). It should be noted that the tests above is just one example of a nonlinearity test. Other possibilities are e.g. the Lagrange Multiplier test discussed previously, see also (Priestley, 1988; Tong, 1990).

The bispectrum and the tests above are useful to check the assumptions on which the methods are build. Furthermore if the stochastic process is nonlinear or non Gaussian, the bispectrum contains valuable information about the process, which can be used for model improvement. In general it is necessary to estimate and analyze higher order moments or spectra in order to be able to establish the need for nonlinear or non-Gaussian components in the model, cf. (Tong, 1990).

5.3 Graphical Methods

In traditional time series modelling a number of graphical methods are used as standard. These includes plot of the time series data, sample autocorrelation function plot, sample partial autocorrelation function plot, sample spectral density functions, histograms, phase diagrams (plotting x_t versus x_{t-1}, etc. These are valuable tools whether the primary interests are linear modeling or not, cf. (Tong, 1990).

A test of the model's ability for simulation is to simulate the system (the deterministic part) with the measured inputs and compare the simulated output with the measured either visually in a plot or by some measure of the distance. This will indicate if the model is able to catch the relevant features of the dynamics of the system, see (Ljung, 1987). If also the statistical properties of the solution is of interest (including the stochastic part of the model), stochastic simulation must be used, cf. Section 2.5. A Monte Carlo technique can be used to simulate a whole family of sample paths of the estimated model. This technique reveals the approximate distribution of the sample paths from the given model.

The estimated parameters are asymptotically normally distributed, cf. Section 3.3.2. A parametric bootstrap technique can then be used to sample a sequence of parameter estimates from the obtained distribution and use each of them in a Monte Carlo simulation, where usually only the deterministic part of the system is simulated. This gives a reasonable illustration of the robustness of the measured variables against variations of the parameters described by the distribution of the estimated parameters, cf. (Holst et al., 1992).

5.4 Cross Validation

The principle of cross validation is to compare the results from a model identified from one data set, $\mathcal{D}_0$, with the same model applied on a new data set, $\mathcal{D}_1$, from the same system. It is expected that the model perform well to the data from which is was identified, but if the same model also perform well to the new data set we believe the model gives a good description of the system.

A formal test for cross validation is the use of the Lagrange Multiplier test described previously in Section 5.1.1. This is done by first estimating the parameters in the model from one data set. With the resulting parameters, the LM test statistic (5.6) is calculated, using the new data set. The hypothesis can be formulated as:

H_0 : The correspondence (as measured by the likelihood function) between the model estimated from $\mathcal{D}_0$ and from $\mathcal{D}_1$ can not be significantly improved by reestimating the parameters from $\mathcal{D}_1$ itself.

H_1 : the opposite statement.

see (Graebe, 1990b). The obtained LM value (5.6) is chi-square distributed with $n = \dim(\theta)$ degrees of freedom. The hypothesis is accepted if $LM < \chi^2_\alpha(n)$ for the chosen level α.

5.5 Bayesian Methods

In this section we will take advantage of the prior information associated with grey box identification in a formal framework. If the prior information can be parameterized via prior probability density functions, which expresses the prior degree of belief, the testing of hypotheses can be handled in a Bayesian framework.

Consider the problem of testing two mutually exclusive and exhaustive hypotheses, (Zellner, 1971):

H_0 : model $\mathcal{M}_0$ with parameters $\theta = \theta_0$

H_1 : model $\mathcal{M}_1$ with parameters $\phi = \phi_1$

where θ_0 and ϕ_1 are specific values of the parameter vectors θ and ϕ, respectively. Our prior probabilities associated with the hypotheses are $p(H_0)$ and $p(H_1)$ with $p(H_0) + p(H_1) = 1$. The prior information is then combined with the sample information to form the posterior probabilities:

$$p(H_0|\mathbf{y}^N) = \frac{p(H_0)\,p(\mathbf{y}^N|\theta = \theta_0)}{p(\mathbf{y}^N)} \tag{5.16}$$

$$p(H_1|\mathbf{y}^N) = \frac{p(H_1)\,p(\mathbf{y}^N|\phi = \phi_1)}{p(\mathbf{y}^N)} \tag{5.17}$$

The *posterior odds* in favor of H_0, denoted by K_{01}, are given by

$$K_{01} = \frac{p(H_0|\mathbf{y}^N)}{p(H_1|\mathbf{y}^N)} = \frac{p(H_0)}{p(H_1)}\frac{p(\mathbf{y}^N|\theta = \theta_0)}{p(\mathbf{y}^N|\phi = \phi_1)} \tag{5.18}$$

From (5.18) it is seen that the posterior odds are the product of the prior odds, $p(H_0)/p(H_1)$, and the likelihood ratio, $p(\mathbf{y}^N|\theta = \theta_0)/p(\mathbf{y}^N|\phi = \phi_1)$. If a decision is to be taken, that is either to accept H_0 or reject H_0, a loss function is needed stating the consequences of the action depending on the truth. Assume that the loss is zero for correct decisions, that is $L(H_0, \hat{H}_0) = L(H_1, \hat{H}_1) = 0$. The loss from selecting H_1 if H_0 is true (error of first kind) is denoted $L(H_0, \hat{H}_1)$ and the loss from selecting H_0 if H_1 is true (error of second kind) is $L(H_0, \hat{H}_1)$. With this loss structure we can evaluate the expected loss associated with the actions accept H_0 and accept H_1 respectively

$$E(L|\hat{H}_0) = p(H_1|\mathbf{y}^N)\, L(H_1, \hat{H}_0) \tag{5.19}$$

$$E(L|\hat{H}_1) = p(H_0|\mathbf{y}^N)\, L(H_0, \hat{H}_1) \tag{5.20}$$

Having calculated the expected losses, we are able to compare them and choose the action which minimizes our expected loss. That is, H_0 is accepted if

$$E(L|\hat{H}_0) < E(L|\hat{H}_1) \tag{5.21}$$

By using (5.16) and (5.17) the inequality can be expressed

$$\frac{p(\mathbf{y}^N|H_0)}{p(\mathbf{y}^N|H_1)} > \frac{p(H_1)\, L(H_1, \hat{H}_0)}{p(H_0)\, L(H_0, \hat{H}_1)} \tag{5.22}$$

In this expression the likelihood ratio, $p(\mathbf{y}^N|H_0)/p(\mathbf{y}^N|H_1)$, is compared with the ratio of prior expected losses. The higher the prior expected loss associated with accepting H_0 in relation to that associated with accepting H_1 the greater the sample evidence in favor of H_0 as reflected by the likelihood ratio on the left hand side of (5.22). This appears to be a sensible procedure for determining the "critical value" in a likelihood ratio test procedure. That is, in a classical likelihood ratio test procedure H_0 is accepted if $p(\mathbf{y}^N|H_0)/p(\mathbf{y}^N|H_1) > \lambda$, where λ is determined by choice of the significance level for the test. In the Bayesian approach *explicit* considerations is given to the loss structure, whereas in the classical test this is done *implicitly* by the choice of significance level.

The test above was a comparison of *simple* hypotheses. In the following the more complicated composite hypotheses are considered (Zellner, 1971)

H_0 : model $\mathcal{M}_0$ with parameters $\theta \sim p(\theta)$

H_1 : model $\mathcal{M}_1$ with parameters $\phi \sim p(\phi)$

where the parameters θ and ϕ are assumed to have the prior probability density functions $p(\theta)$ and $p(\phi)$, respectively. This is opposed to the situation above, where only specific values of the parameters was considered. The posterior probability for H_0 can be expressed

$$p(H_0|\mathbf{y}^N) = \frac{p(H_0)\int p(\theta)p(\mathbf{y}^N|\theta)\,d\theta}{p(\mathbf{y}^N)} \tag{5.23}$$

From this expression and a similar one for $p(H_1|\mathbf{y}^N)$, we can calculate the posterior odds in favor of H_0

$$K_{01} = \frac{p(H_0|\mathbf{y}^N)}{p(H_1|\mathbf{y}^N)} = \frac{p(H_0)}{p(H_1)}\frac{\int p(\theta)p(\mathbf{y}^N|\theta)\,d\theta}{\int p(\phi)p(\mathbf{y}^N|\phi)\,d\phi} \tag{5.24}$$

As pointed out by Zellner (1971) the posterior odds are equal to the prior odds times the ratio of *averaged* likelihoods with the prior pdf's $p(\theta)$ and $p(\phi)$ serving as weighting functions. This contrasts with the usual likelihood ratio testing procedure which involves taking the ratio of *maximized* likelihood functions under H_0 and H_1. If a decision is to be made on acceptance or rejection of H_0 then a loss function is needed. An inequality similar to (5.22) can be made for minimizing the expected loss, except that the ratio of maximized likelihood functions is replaced by the ratio of averaged likelihood functions weighted by the prior pdf's of the parameters under H_0 and H_1, respectively.

5.6 Summary

In this chapter a number of different methods have been discussed which can used in validation by comparing and testing particular models. Some of the methods also have diagnostic properties which is useful for indicating the next step to take in the iterations of building the model. For model validation it is useful to have access to a number of different methods and tests, but normally just a few of them are necessary for an initial validation of a model. The following constitute a rough list of simple (and fast) methods, sufficient for this initial validation:

- Test the parameters of the model for significance and check the correlation matrix of the parameters for near singularities.
- Test the residuals for whiteness and for independence of the inputs.
- Compare measured and simulated outputs.
- Compare estimated physical parameters with prior information.

If the model is accepted in these tests, it should be followed by some cross validation. This could be done simply by splitting the available data set in two parts and estimating the parameters from the one part and testing the resulting model against the other part of the data set. Dependent of the intended use of the model some of the more powerful (and cumbersome) tests may be used.

The advantage of grey box modelling is the available prior information. This information can be incorporated in a unified way when comparing and testing hypothesis by using Bayesian methods. The Bayesian decision theoretic approach for testing (and choosing between) simple hypotheses leads to a test procedure similar to the likelihood ratio test, except that the critical value in the test is not determined from an arbitrary chosen significance level, but rather from explicit consideration of the loss structure and prior information.

The choice of level for acceptance will always be the subjective decision of the user. In this context the intended use of the model has a large influence. Hence, as stressed in the beginning of the chapter, it is preferable that the class of purposive models is defined by the experimental conditions under which the model is identified. This approach simplifies the model validation. If, however, a class of purposive models can be defined, it may be adequate to evaluate the resulting models using a Bayesian decision theoretic approach and define the purposivity via loss functions, see also (Holst et al., 1992).

Chapter 6

Experimental Design

The purpose of experiment design is to ensure that the information provided by the experiment is maximized within given limitations. The design of experiments for identification of dynamical models includes choice of input and output ports of the system, as well as choice of input signals, sampling time, presampling filters and, if necessary, filtering of input and output data.

Any design must take into account the constraints on the allowable experimental conditions. Some typical constraints that might be met in practice are amplitude constraints on inputs, outputs or internal signals; power constraints on signals; limited total experimental time or limited number of samples. Further more, the available measureing and/or actuating equipment introduces constraints like least possible sampling interval, highest possible accuracy, etc.

There is a substantial litterature on optimal experiment design. In the books by Fedorov (1972) and Pázman (1986), among others, the theory and mathematical treatment of optimal experiment design in general is considered. Surveys on this subject specifically for dynamic systems identification are given by Mehra (1974), and more recently by Walter & Pronzato (1990). The books by Goodwin & Payne (1977) and Zarrop (1979) are important contributions for optimal experiment design for linear dynamic systems. Design of experiments for identification of transfer functions in the frequency domain has been considered by e.g. Yuan & Ljung (1985), who discuss the use of prior information. The intended use of the transfer function models is considered by Gevers & Ljung (1986).

Prior information about the system and the intended use of the model are both inevitable terms in connection with design of experiments. Concerning physical models there are certain characteristics which should be reflected in the design. The criteria of optimality should be formulated in terms of the interesting physical characteristics of the system. Especially for physical models there are often available prior information about the model. This prior information should also be used in the design of experiments. This can be done either in a classical way using only prior expected values of the parameters or in a Baysian approach, by incorporating prior distributions of the parameters for the design of optimal experiments, see (Melgaard, Sadegh, Madsen, & Holst, 1993; Sadegh, Melgaard, Madsen, & Holst, 1994).

6.1 Structural Identifiability

It is a crucial question whether the parameters of a specified model can be identified. If a non identifiable model is specified, the methods for estimation will not converge. Identifiability is a concept that addresses the problem whether the given identification procedure will yield a unique value of the parameters. One aspect of the problem has to do with the experimental conditions, e.g. whether the data set is informative enough (persistently exciting) to distinguish between different models for a given estimation method. Another aspect has to do with the model structure, i.e. if different sets of parameters will give equal models, given that the data is informative. This qualitative aspect of experiment design has to do with the selection of input and output ports for the model, the parameterization and structure of the model. Hence, the questions of structural identifiability can be answered even before the data have been collected on the system. The parameter θ_i is said to be structurally globally identifiable if for almost any θ^*

$$M(\theta) = M(\theta^*) \Rightarrow \theta_i = \theta_i^*. \tag{6.1}$$

It is called structurally locally identifiable if for almost any θ^* there exits a neighborhood $\upsilon(\theta^*)$ such that if $\theta \in \upsilon(\theta^*)$, then the implication (6.1) is true.

For convenience the input, here, is assumed to be generated independently from the output. The inclusion of a feedback may result in loss of identifiability, see e.g. the example given in Chap. 14 by Ljung (1987). For none of the cases in this study it was necessary to consider feedback. It may be necessary, though, in other cases to operate under closed-loop control. This can be for safety reasons if e.g. the system is unstable. If there are limitations on the maximum allowable power of the output signals it may even be advantageous to use feedback. In these cases it is possible to retain identifiability by considering the feedback mechanism properly, see (Söderström & Stoica, 1989, Chap. 10).

The notion of structural identifiability is of importance for both linear and nonlinear models, but there are some specific properties which differs for the two types. In this context it is relevant to speak of two kinds of nonlinearities of the model, one is nonlinear *in the inputs*, which means that the output does not satisfy the superposition principle concerning the inputs:

$$\mathbf{y}_{\mathcal{M}}(\theta, \lambda \mathbf{u}_1 + \mu \mathbf{u}_2, t) \neq \lambda \mathbf{y}_{\mathcal{M}}(\theta, \mathbf{u}_1, t) + \mu \mathbf{y}_{\mathcal{M}}(\theta, \mathbf{u}_2, t) \tag{6.2}$$

for some real λ and μ. The other kind of nonlinearity, is when the model is nonlinear *in the parameters*, which means that the output from the model does not satisfy the superposition principle concerning the parameters:

$$\mathbf{y}_{\mathcal{M}}(\lambda \theta_1 + \mu \theta_2, \mathbf{u}, t) \neq \lambda \mathbf{y}_{\mathcal{M}}(\theta_1, \mathbf{u}, t) + \mu \mathbf{y}_{\mathcal{M}}(\theta_2, \mathbf{u}, t) \tag{6.3}$$

Most often phenomenological models are nonlinear in the parameters.

Remark 6.1 *It should be remarked that only models that are nonlinear in the parameters can have structurally locally identifiable parameters which are not at the same time structurally globally identifiable, see (Walter & Pronzato, 1990). The outputs of models that are linear in the parameters can be written as*

$$\mathbf{y}_{\mathcal{M}}(t, \theta, \mathbf{u}) = \phi(t, \mathbf{u})'\theta \tag{6.4}$$

and

$$M(\theta) = M(\theta^*) \Leftrightarrow \phi(t, \mathbf{u})'\theta = \phi(t, \mathbf{u})'\theta^* \tag{6.5}$$

This set of linear equations has a unique solution if the columns of ϕ *are linearly independent, or else it has an infinite number of solutions.*

Remark 6.2 *Ljung & Glad (1994) have shown that testing global identifiability is equivalent to the possibility to express the model structure as a linear regression*

$$\psi(t) = \phi(t)'\theta \, . \tag{6.6}$$

This is a generalization of (6.4) where $\psi(t)$ *and* $\phi(t)$ *are matrices, entirely formed from past values of* $\mathbf{y}(t)$, $\mathbf{u}(t)$ *and their derivatives. Ljung & Glad (1994) treat this problem for models that have polynomial nonlinearities with respect to the parameters and the inputs.*

In the following when the model is only refered to as linear or nonlinear this means with respect to the inputs.

6.1.1 Linear Models

The problem of structural identifiability of linear state space models arises from the fact that for a given transfer function model corresponds, in general, a continuum of state space models. It is therefore necessary to put restrictions on the structure of the state space model in order to provide a unique relation between the parameters of the state space model and the transfer function.

In this section the descriptions are kept in continuous time even though our observations are in discrete time for any practical applications. It is obvious that the sampling of the observations (inputs and outputs) may deteriorate the identifiability properties of the model, but concerning only the structural identifiability it will have no influence.

Consider a continuous time state space model, which is linear in the states but in general nonlinear in the parameters

$$d\mathbf{x}(t) = \mathbf{A}(\theta)\mathbf{x}(t)\,dt + \mathbf{B}(\theta)\mathbf{u}(t)\,dt + d\mathbf{e}_1(t), \tag{6.7}$$

$$\mathbf{y}(t)\,dt = \mathbf{C}(\theta)\mathbf{x}(t)\,dt + \mathbf{D}(\theta)\mathbf{u}(t)\,dt + d\mathbf{e}_2(t), \tag{6.8}$$

where the state vector $\mathbf{x}(t)$ has dimension n, the input $\mathbf{u}(t)$ has dimension m and the output $\mathbf{y}(t)$ has dimension s, $d\mathbf{e}_1(t)$ and $d\mathbf{e}_2(t)$ are mutually independent Wiener processes with incremental variances $\mathbf{\Sigma}_1(\theta)\,dt$ and $\mathbf{\Sigma}_2(\theta)\,dt$ of dimension n and s respectively. The part of the output, that cannot be predicted from past data is $d\epsilon(t) = (\mathbf{y}(t) - \hat{\mathbf{y}}(t))\,dt$. This quantity, denoted by $d\epsilon(t)$ is called the innovation. It is a Wiener process of dimension s with incremental variance $\Sigma(\theta)\,dt$. The incremental variance of the innovation process $d\epsilon(t)$ is calculated from

$$\mathbf{\Sigma}(\theta) = \mathbf{C}(\theta)\bar{\mathbf{P}}(\theta)\mathbf{C}(\theta)^T + \mathbf{\Sigma}_2(\theta) \tag{6.9}$$

where $\bar{\mathbf{P}}(\theta)$ is the positive semidefinite solution to the stationary Riccati equation:

$$\begin{aligned} 0 = {} & \mathbf{A}(\theta)\bar{\mathbf{P}}(\theta) + \bar{\mathbf{P}}(\theta)\mathbf{A}(\theta)^T + \mathbf{\Sigma}_1(\theta) \\ & -\bar{\mathbf{P}}(\theta)\mathbf{C}(\theta)^T\mathbf{\Sigma}_2(\theta)^{-1}\mathbf{C}(\theta)\bar{\mathbf{P}}(\theta). \end{aligned} \tag{6.10}$$

By rewriting (6.7) and (6.8) we obtain the *innovations form* of the state space model

$$d\hat{\mathbf{x}}(t) = \mathbf{A}(\theta)\hat{\mathbf{x}}(t)\,dt + \mathbf{B}(\theta)\mathbf{u}(t)\,dt + \mathbf{K}(\theta)d\epsilon(t) \tag{6.11}$$

$$y(t)\,dt = \mathbf{C}(\theta)\hat{\mathbf{x}}(t)\,dt + \mathbf{D}(\theta)\mathbf{u}(t)\,dt + d\epsilon(t) \tag{6.12}$$

where $\mathbf{K}(\theta)$, the continuous time Kalman gain is given as

$$\mathbf{K}(\theta) = \bar{\mathbf{P}}(\theta)\mathbf{C}(\theta)^T\mathbf{\Sigma}_2(\theta)^{-1}. \tag{6.13}$$

Introducing p as the differentiation operator the corresponding transfer function model has the form

$$y(t)\,dt = \mathbf{G}(p,\theta)\mathbf{u}(t)\,dt + \mathbf{H}(p,\theta)d\epsilon(t), \tag{6.14}$$

where the transfer functions are determined as

$$\mathbf{G}(p,\theta) = \mathbf{C}(\theta)(p\mathbf{I} - \mathbf{A}(\theta))^{-1}\mathbf{B}(\theta) + \mathbf{D}(\theta), \tag{6.15}$$

$$\mathbf{H}(p,\theta) = \mathbf{C}(\theta)(p\mathbf{I} - \mathbf{A}(\theta))^{-1}\mathbf{K}(\theta) + \mathbf{I}. \tag{6.16}$$

From the observations of $\mathbf{u}(t)$ and $y(t)$ we are able to observe the transfer matrices $\mathbf{G}(p)$ and $\mathbf{H}(p)$ with elements of the form

$$\mathbf{G}_{ij}(p) = \frac{b_0p^\gamma + b_1p^{\gamma-1} + \cdots + b_\gamma}{p^\delta + a_1p^{\delta-1} + \cdots + a_\delta} \tag{6.17}$$

By comparing the parameterization of (6.16) with elements of the form (6.17) it is possible to check the structural identifiability of the parameters of a specified model. Then we have

$$\begin{aligned} &M(\theta) = M(\theta^*) \Leftrightarrow \\ &\qquad \mathbf{G}(p,\theta) = \mathbf{G}(p,\theta^*) \wedge \mathbf{H}(p,\theta) = \mathbf{H}(p,\theta^*), \end{aligned} \tag{6.18}$$

and the identifiability properties of the model are directly related to the number of solutions for θ of this equation.

There are different ways of parameterizing the noise model of the system. One possibility is directly to estimate the stationary Kalman gain.

It is easy to verify the identifiability of this parameterization of the noise model, through (6.16), but it is on the other hand difficult to use physical insight about the noise with this parameterization. The other way to specify the noise model, is in terms of the incremental variance of the process noise, $\mathbf{\Sigma}_1(\theta)\,dt$ and measurement noise $\mathbf{\Sigma}_2(\theta)\,dt$. Verifying structural identifiability is more complicated in this case, because the solution of the stationary Riccati equation (6.10) is needed for calculation of the Kalman gain (6.13) to be inserted in (6.16). Generally it is not possible to estimate all elements of the variance matrices, but e.g. only the diagonals. This is due to the fact that the direct parameterized innovations form contain $n_{dir} = ns + s(s+1)/2$ elements for the noise properties, through $\mathbf{K}$ and $\mathbf{\Sigma}$, whereas the model with separated noise sources has $n_{sep} = n(n+1)/2 + s(s+1)/2$ elements, assuming the process noise and measurement noise are mutually independent. This means, that if a model, with its noise structure specified through the direct innovations form, is identifiable, then n_{dir} is the maximum number of identifiable "noise"-parameters. By noise-parameters is meant the subset of θ which is not identifiable via $\mathbf{G}(p,\theta)$. It is, on the other hand, easier to implement the physical insight by a separated noise structure, for instance that the noise is only affecting certain states of the model.

Example 6.1 *The problem of structural identifiability of linear timeinvariant models is illustrated by an example. Consider a model (6.7) and (6.8) with matrices:*

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} -a & a \\ b & -(b+c) \end{bmatrix}, & \mathbf{B} &= \begin{bmatrix} 0 \\ d \end{bmatrix}, \\ \mathbf{C} &= \begin{bmatrix} 0 & 1 \end{bmatrix}, & \mathbf{D} &= \begin{bmatrix} 0 \end{bmatrix} \end{aligned} \tag{6.19}$$

with parameters $\theta = (a\ b\ c\ d)'$. *The transfer function* $G(p)$ *as given* *by (6.15) is*

$$G(p) = \frac{d(p+a)}{p^2 + (a+b+c)p + ac} \tag{6.20}$$

This must be compared with the structure we are able to observe, namely

$$G(p) = \frac{b_0 p + b_1}{p^2 + a_1 p + a_2} \tag{6.21}$$

By comparing (6.20) and (6.21), it is seen that we are able to identify a, b, c *and* d *from* b_0, b_1, a_1 *and* a_2*. Hence the model is structural identifiable. For the noise model* $H(p)$ *given by (6.16), there a more possibilities for parameterization. One possibility is a direct parameterization of the innovations form, i.e. directly estimation of the kalman gain*

$$\mathbf{K} = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} \tag{6.22}$$

By calculating the noise model, $H(p)$*, from (6.16) we optain:*

$$H(p) = \frac{p^2 + (a+b+c+k_2)p + (k_1 b + k_2 a + ac)}{p^2 + (a+b+c)p + ac}. \tag{6.23}$$

Comparing this model with the modelstructure (6.21) it is seen, that it is possible to identify k_1, k_2 *and the variance of the innovations* $\boldsymbol{\Sigma}$*, when already the parameters of* $G(p)$ *in (6.20) are identifiable. The total model including the noise model is in this case structural identifiable. Another way to parameterize the noise model is via the matrices* $\boldsymbol{\Sigma}_1$ *and* $\boldsymbol{\Sigma}_2$*. These matrices are connected to* $\mathbf{K}$ *and* $\boldsymbol{\Sigma}$ *through (6.13),* *(6.10) and (6.9) in a complex way. From (6.9) we have*

$$\Sigma = P_{22} + \Sigma_2, \tag{6.24}$$

and from (6.13)

$$K = \begin{bmatrix} P_{12}/\Sigma_2 \\ P_{22}/\Sigma_2 \end{bmatrix}. \tag{6.25}$$

Since, in this case, we are able to identify Σ and K, then it is also possible to identify P_{12}, P_{22} and Σ_2. From the stationary Riccati equation (6.10) a set of equations must be solved

$$0 = -2aP_{11} + 2aP_{12} + \Sigma_{1,11} - P_{12}^2/\Sigma_2 \tag{6.26}$$

$$0 = 2bP_{12} - 2(b+c)P_{22} + \Sigma_{1,22} - P_{22}^2/\Sigma_2 \tag{6.27}$$

$$0 = bP_{11} - (a+b+c)P_{12} + aP_{22} + \Sigma_{1,12} - P_{12}P_{22}/\Sigma_2 \tag{6.28}$$

From (6.26) it is seen that a linear combination of P_{11} and $\Sigma_{1,11}$ is identifiable. From (6.27) $\Sigma_{1,22}$ can be identified. A linear combination of P_{11} and $\Sigma_{1,12}$ is identified from (6.28). Since both $\mathbf{P}$ and $\mathbf{\Sigma}_1$ have to be positive semidefinite it is not possible to identify the full parameterized matrices. By fixing $\Sigma_{1,12} = 0$ all the other parameters are identifiable. This problem could already be seen by considering the dimensions of the matrices. The dimensions of K is $n \times s$ and $\mathbf{\Sigma}$ has $s(s+1)/2$ elements, where n is the order of the system and s is number of outputs from the system, $n_{dir} = ns + s(s+1)/2$ is the maximum number of parameters which can be identified for $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$. In this example we are able to identify $n_{dir} = 3$ parameters for the noise model, as previously seen. This is done by considering only the diagonals of $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$, thus fixing $\Sigma_{1,12} = 0$. □

6.1.2 Nonlinear Models

Testing for structurally identifiability in a nonlinear model is slightly more complicated. One method consist of linearizing the model around some equilibrium point and then apply the method for linear models. A second approach uses a series expansion of the output either in the time domain or in the time and input domain (Walter & Pronzato, 1990). $M(\theta) = M(\theta^*)$ then implies that the coefficients of the series should be equal, which, as in the linear case, yields a set of equations in θ parameterized by θ^*.

6.2 Criteria for Optimality

When the model is made structurally identifiable by proper selection of input and output ports and model structure, we still have to select the best experiment to be performed for collecting maximum information of the system to be identified, with respect to the selected method of identification.

In order to perform the quantitative part of the experiment design, a measure of the information achieved from an experiment is needed. It is common practice, cf. (Goodwin & Payne, 1977), to select a performance measure related to the expected accuracy of the parameter estimates to be obtained, in general the covariance of the parameters. The Cramér-Rao inequality, see Theorem 3.3, gives a limit to the covariance of any unbiased estimator $\hat{\theta}(\mathbf{Y})$ of θ, subject to certain regularity conditions, cf. (Rao, 1973),

$$V(\hat{\theta}) \geq \mathbf{M}_F^{-1} \tag{6.29}$$

where $\mathbf{M}_F$ is the *Fisher information matrix*, defined by

$$M_F = E_{Y|\theta}\left\{ \left(\frac{\partial \log p(\mathbf{Y}|\theta)}{\partial\theta}\right)\left(\frac{\partial \log p(\mathbf{Y}|\theta)}{\partial\theta}\right)^T \right\}, \tag{6.30}$$

$\mathbf{Y}$ is a vector of all observations, θ is the unknown parameter vector, and $p(\mathbf{Y}|\theta)$ is the conditional probability density of $\mathbf{Y}$ for given θ. When the estimator is asymptotically efficient, the rationale for using the Fisher information matrix as a suitable characterization of the asymptotic parameter uncertainty is obtained.

It should be remarked that an experiment satisfying $\det \mathbf{M}_F(\theta) \neq 0$ is called informative, this ensures local identifiability for the model parameters. Designing an experiment by minimizing a suitable criterion

$$J(\xi) = \phi(\mathbf{M}_F(\theta, \xi)) \tag{6.31}$$

can thus be seen as maximizing a measure of identifiability, where ϕ is a scalar function and ξ is the design.

6.2.1 Local Design

A number of different standard measures of information has been studied. For a review of the properties of the different measures of information see (Pázman, 1986; Walter & Pronzato, 1990). A large number of these standard measures are particular cases of the general L_k-class of the optimality criteria. The criterion belonging to this class is defined by a function of the form

$$\phi(\mathbf{M}_F) = \begin{cases} [n^{-1}\mathrm{tr}(\mathbf{V}\mathbf{M}_F^{-1}\mathbf{V}^T)^k]^{1/k} & \text{if } \det \mathbf{M}_F \neq 0 \\ \infty & \text{if } \det \mathbf{M}_F = 0 \end{cases} \tag{6.32}$$

where $k > 0$ and $\mathbf{V}$ is a nonsingular $n \times n$ matrix. The local optimal design refers to the case where the criterion is minimized for a given value of the parameters θ^*, say. This value is often chosen as the expected mean $\theta^* = E_\theta(\theta)$ calculated from the prior distribution of the parameters $p(\theta)$.

D-optimality

The most studied criterion is the *D-optimality* criterion, which is defined by

$$\phi(\mathbf{M}_F(\xi)) = -\log\det \mathbf{M}_F(\xi) \; . \tag{6.33}$$

The criterion is obtained for $\mathbf{V} = \mathbf{I}$ and $k \to 0$ in (6.32). Thus this design minimizes the generalized variance of the parameter estimates. The criterion has a geometrical interpretation: The asymptotic confidence regions for the maximum likelihood estimate of θ are ellipsoids, and a D-optimal experiment thus minimizes the volume of these ellipsoids. An important

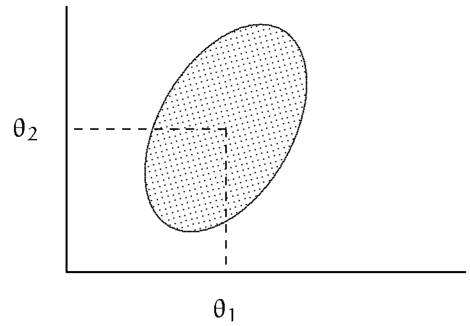


Figure 6.1. *D-optimal design minimizes the volume of the confidence ellipsoid.*

property of D-optimal experiments is their independence of the scaling of parameters, due to the geometrical property of the determinant.

L-optimality

Criteria that are linear in the inverse Fisher information matrix (= dispersion matrix of the best linear estimator) are obtained from (6.32) by setting $k = 1$. The particular choices $\mathbf{V} = \mathbf{I}$ and $\mathbf{V} = \text{diag}(\theta_i^{-1}, i = 1, \cdots, n)$ respectively, correspond to *A*- and *C-optimality*. A-optimal experiments minimize the sum of variances of θ. C-optimal experiments are related to the relative precision of the estimates, actually they summarize $1/t_i^2$, where t_i is the t-statistic of the ith parameter. This criterion is also independent of the parameter scale. The *A*- and *C-optimality* criteria does not take

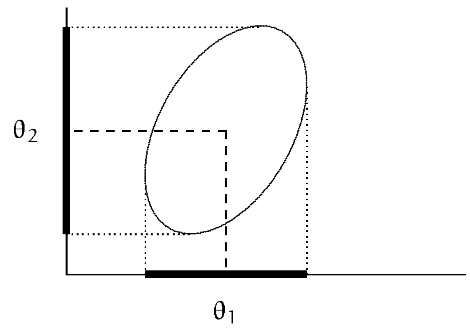


Figure 6.2. *C-optimal design minimizes the relative precision of the single parameters, without paying attention to the correlation structure.*

the correlation between the parameters into account. From a pragmatic point of view such criteria should be seen with suspicion because they can lead to design of experiments in which the parameters are unidentifiable, $\det \mathbf{M}_F = 0$, see (Goodwin & Payne, 1977).

E-optimality

The *E-optimality* criterion is obtained from (6.32) by setting $\mathbf{V} = \mathbf{I}$ and $k \rightarrow \infty$, this corresponds to minimizing the maximum eigenvalue of $\mathbf{M}_F^{-1}$,

$$\phi(\mathbf{M}_F(\xi)) = \lambda_{max}(\mathbf{M}_F(\xi)^{-1}) \tag{6.34}$$

Geometrically this can be understood as minimizing the maximum diameter of the asymptotic confidence ellipsoids for the parameters, because the semi-axes of the ellipsoids are directed as the eigenvectors of $\mathbf{M}_F^{-1}$ with lengths proportional to the eigenvalues of the matrix. In other words, an E-optimal design aims at improving the most uncertain region of the parameter space and making the confidence region as spherical as possible. By using $\mathbf{V} = \text{diag}(\theta_i^{-1}, i = 1, \cdots, n)$, the criterion is independent of the

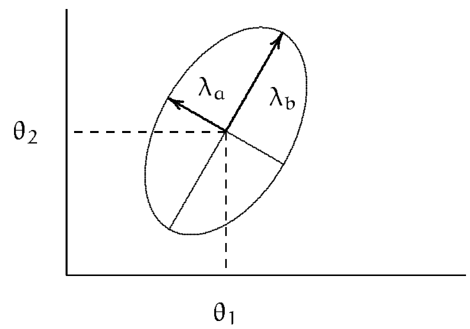


Figure 6.3. *E-optimal design minimizes the maximum eigenvalue of the inverted information matrix, thus making the confidence region of the parameters as spherical as possible.*

scaling of parameters. The geometrical interpretation is still valid, but for the *normalized* inverted information matrix.

6.2.2 Physical Measures

Other criteria for optimality than the standard ones can be specified. Specifically when dealing with physical models nonstandard criteria may be of interest, since the physical parameters of greatest interest might not be directly entering into the system description but rather through some transformation of these parameters. Assume that the parameters of interest are described by the functional relation $\mathbf{f}(\theta)$, where θ is the parameters of the model. It is important to take this transformation into account in the design of experiments. The asymptotic covariance of $\mathbf{f}$ is calculated by using Gauss' formula

$$\mathbf{M}_{F,f}^{-1} = \frac{\partial \mathbf{f}}{\partial \theta^T} \mathbf{M}_F^{-1}(\theta) \frac{\partial \mathbf{f}}{\partial \theta} . \tag{6.35}$$

Any of the standard measures of information can now be applied for $\mathbf{M}_{F,f}$. In Melgaard, Madsen, & Holst (1992b) a physical model of the thermal characteristics of a building is considered. The classical measures of information all lead to the same optimal design of input signal, a certain PRBS sequence. It turned out, however, that the main interest is not focused on the individual parameters of the model, but on the overall heat transmittance and internal heat capacity of the building. These physical characteristics are given as a function of the model parameters. An optimal design considering this application specific measure of information leads to an input signal, which is a step. This design turns out to be much different from the original design, (more weight on low frequencies), but it reflects the demands of the building physicists.

6.2.3 Bayesian Design

In general the designs of the previous sections are dependent upon the unknown parameter values. Any design will and shall be dependent upon the prior information available about the system to be identified. From a statistical point of view it is obvious, that one way of incorporating this prior information is to use a Bayesian procedure. In this approach the prior knowledge of the parameters are expressed via their the prior distribution function, which expresses the partial information on the system available prior to an experiment.

In the Bayesian approach a loss function $L(\theta, \hat{\theta})$ is defined, which describes the consequences of obtaining $\hat{\theta}$, when θ is the true parameter vector. This function is then used as a basis for obtaining the optimal experiment design. Prior to the experiment the expected performance is

$$J_1 = E_{\theta,Y}[L(\theta, \hat{\theta})]. \tag{6.36}$$

The criterion may be optimized directly with respect to the allowable experimental conditions. By making suitable assumptions and using truncated Taylor expansion J_1 can be simplified to:

$$\begin{aligned} J_1 &= E_{\theta,Y}[L(\theta, \hat{\theta})] = E_\theta E_{Y|\theta}[L(\theta, \hat{\theta})] \\ &\simeq E_\theta[L(\theta, \theta) + \frac{1}{2}\mathrm{tr}\{(\partial^2 L/\partial\hat{\theta}^2)\boldsymbol{M}_F^{-1}\}] , \end{aligned} \tag{6.37}$$

where $\boldsymbol{M}_F$ is given by (6.30) and it is assumed that $\hat{\theta}$ is an efficient unbiased estimator. Optimizing this criterion is thus equivalent to optimizing a criterion of the form $E_\theta(\mathrm{tr}\{\boldsymbol{W}\boldsymbol{M}_F^{-1}\})$, where $\boldsymbol{W} = \partial^2 L/\partial\hat{\theta}^2$ is a weight matrix, see (Goodwin & Payne, 1977). Generalizing this criterion yields

$$J_2 = E_\theta(\phi(\boldsymbol{M}_F(\theta, \xi))) , \tag{6.38}$$

where expectation is taken over the prior distribution of θ, and ϕ is suitable scalar function. In general ϕ can be of the form considered in the previous sections for local designs. As pointed out by Walter & Pronzato (1990) there are several average optimality criteria (6.38) corresponding to one local design (6.31). This is due to the fact that integration is *not* a commutative operator with nonlinear functions, e.g. inversion and logarithm. This implies that considering the maximization of $\phi = \det \mathbf{M}_F$ leads to a different design than considering the minimization of $\phi = (\det \mathbf{M}_F)^{-1}$ in the criterion (6.38). For a local design these functions result in the same design.

6.2.4 Minimax Design

The Bayesian design presented in the previous section is a design criterion based on the expected prior performance. The criterion takes into account the whole prior distribution of the parameters. This approach is preferred compared to the classical case of local design, where the prior distribution, and specifically the uncertainty, of the parameters is not considered. However optimization of this criterion does not give any guard against a single non-informative experiment, because the average performance does not highlight the worst case performance. One might prefer to maximize the worst possible performance of the designed experiment. Such *minimax* approach relies on the knowledge of some prior admissible domain of the parameters $\theta \in \mathbf{\Theta}$, without requesting any prior information about θ. Minimax criteria can be deduced from local optimality criteria by

$$J(\xi) = \max_{\theta \in \Theta} \phi(\mathbf{M}_F(\theta, \xi)) , \tag{6.39}$$

where $J(\xi)$ is to be minimized with respect to ξ. Classical functions ϕ correspond to L- and D-optimality, see (Walter & Pronzato, 1990).

6.3 Design of Optimal Inputs

We will now evaluate the information matrix in order to determine optimal designs of input signals. Consider for simplicity a scalar discrete time model of the form:

$$y_t = G_1(q)u_t + G_2(q)\epsilon_t \tag{6.40}$$

where $\{u_t\}$ and $\{y_t\}$ are the input and output sequences, respectively, and $\{\epsilon_t\}$ is a sequence of Gaussian random variables with variance σ^2. G_1 and G_2 are transfer functions in the shift-operator. In the following an expression for M_F is derived for the model (6.40). Define ϵ_t as the residual sequence, then:

$$\epsilon_t = G_2^{-1}(q)[y_t - G_1(q)u_t] \; . \tag{6.41}$$

By using this expression in the definition of Fisher's information matrix (6.30), we obtain

$$M_F = E_{Y|\theta}\{\frac{1}{\sigma}\sum_{t=1}^{N}(\frac{\partial \epsilon_t}{\partial \theta})(\frac{\partial \epsilon_t}{\partial \theta})^T\} + \frac{N}{2\sigma^2}(\frac{\partial \sigma}{\partial \theta})(\frac{\partial \sigma}{\partial \theta})^T \; , \tag{6.42}$$

Goodwin & Payne (1977, p. 131). Differentiating expression (6.41) and using superposition yields:

$$\frac{\partial \epsilon_t}{\partial \theta} = (\frac{\partial \bar{\epsilon}_t}{\partial \theta}) + (\frac{\partial \tilde{\epsilon}_t}{\partial \theta}) \; , \tag{6.43}$$

where

$$\frac{\partial \bar{\epsilon}_t}{\partial \theta} = -G_2^{-1}(q)(\frac{\partial G_1(q)}{\partial \theta})u_t \ , \qquad (6.44)$$

$$\frac{\partial \tilde{\epsilon}_t}{\partial \theta} = -G_2^{-1}(q)(\frac{\partial G_2(q)}{\partial \theta})\epsilon_t \ . \qquad (6.45)$$

After substituting these expressions in (6.42) and assuming no feedback in the system ($\{u_t\}$ is independent of $\{\epsilon_t\}$), we get:

$$M_F = \frac{1}{\sigma}\sum_{t=1}^{N}(\frac{\partial \bar{\epsilon}_t}{\partial \theta})(\frac{\partial \bar{\epsilon}_t}{\partial \theta})^T + M_c \ , \qquad (6.46)$$

where M_c does not depend upon the choice of input signal.

Now make the following simplifying assumptions:
1) The experiment time (i.e. N) is large.
2) The input $\{u_t\}$ is restricted to the class admitting a spectral representation with spectral distribution function $F(\omega)$, $\omega \epsilon [-\pi, \pi]$.
3) The allowable input power is constrained.

Since N is large it is more convenient to work with the average information matrix, which gives the following:

$$\bar{M}_F = \lim_{N \to \infty} \frac{1}{N} M_F = \frac{1}{\pi}\int_0^{\pi} (\tilde{M}_F(\omega) + \bar{M}_c)\, d\xi(\omega) \ , \qquad (6.47)$$

where $\xi(\omega)$ is defined by

$$d\xi(\omega) = \begin{cases} \frac{1}{2} dF(\omega) & \omega = 0, \omega = \pi \\ dF(\omega) & \omega \epsilon]0, \pi[\end{cases}$$

and

$$\tilde{M}_F(\omega) = \mathrm{Re}\{\frac{1}{\sigma}[\frac{\partial G_1(e^{j\omega})}{\partial \theta}]G_2^{-1}(e^{j\omega})$$

$$\times G_2^{-1}(e^{-j\omega})[\frac{\partial G_1(e^{-j\omega})}{\partial\theta}]^T\} \tag{6.48}$$

and

$$\begin{aligned}\bar{M}_c(\omega) = \frac{1}{2\pi}\int_{-\pi}^{\pi}&[\frac{\partial G_2(e^{j\omega})}{\partial\theta}]G_2^{-1}(e^{j\omega}) \\ &\times G_2^{-1}(e^{-j\omega})[\frac{\partial G_2(e^{-j\omega})}{\partial\theta}]^T d\omega \\ &+\frac{1}{2\sigma^2}(\frac{\partial\sigma}{\partial\theta})(\frac{\partial\sigma}{\partial\theta})^T .\end{aligned} \tag{6.49}$$

It should be remarked that both $\tilde{M}_F(\omega)$ and $\bar{M}_c(\omega)$ are dependent upon the parameters θ. This means that if a local design is used then only $\tilde{M}_F(\omega)$ need to be considered. On the other hand if a Bayesian design like (6.38) is used then both $\tilde{M}_F(\omega)$ and $\bar{M}_c(\omega)$ must ve evaluated.

The power restriction of the input signal can be formulated as

$$P_u = \frac{1}{\pi}\int_0^{\pi} d\xi(\omega) = 1 . \tag{6.50}$$

We are now ready to give the following theorem, which states that it is always possible to find an optimal input comprising a finite number of sinusoids.

Theorem 6.1 *For any power constrained design $\xi_1(\omega)$ with corresponding $n \times n$ average information matrix $\bar{M}_F(\xi_1)$, there always exists a power constrained design $\xi_2(\omega)$ which is piecewise constant with at most $n(n+1)/2+1$ discontinuities and $\bar{M}_F(\xi_2) = \bar{M}_F(\xi_1)$. For the design criterion $J = -\log\det\bar{M}_F$, optimal designs exist comprising not more than $n(n+1)/2$ sinusoids.*

PROOF: *Only an outline of the proof is given. It can be shown that the set of all average information matrices corresponding to input power* *constrained designs is the convex hull of the set of all average information matrices corresponding to single frequency designs. Hence, from Caratheodory's theorem (see (Fedorov, 1972)) the result for the first part of the theorem follows. For a convex function of the information matrix the optimal design is a boundary point of the convex hull, therefore one less sinusoidal component is needed. For the complete proofs see (Fedorov, 1972; Goodwin & Payne, 1977).* □

6.3.1 Bayesian Approach

We now turn to the Bayesian approach, considering the generalized criterion (6.38). Assume that the prior knowledge of the system is given in terms of a prior distribution of the parameters. Hence we are able to evaluate the expectation of the considered criterion with respect to the prior distribution of the parameters cf. (6.38), in stead of simply evaluating the criterion at some fixed values of the parameters. In general the resulting design will be different for the two approaches, cf. the example following. In the Bayesian case it is possible to prove a theorem similar to Theorem 6.1.

Theorem 6.2 *Using* $J = E_\theta(-\log\det\bar{M})$ *as criterion, optimal designs exist comprising not more than* $n(n+1)/2$ *sinusoidal components.*

PROOF: *The criterion may be written as*

$$J = \int_{\Omega_0} -\log\det\bar{M}(\xi(\omega),\theta))\, p(\theta)\, d\theta \ , \tag{6.51}$$

where $\Omega_\theta \subset \mathbb{R}^n$ *is assumed to be a closed and bounded interval of the parameters and* $p(\theta)$ *is the prior probability density of the parameters.*

The mean-value theorem states, that for all $\xi(\omega)$ there exists a $\theta^* \in \Omega_\theta$ such that

$$J = K(-\log\det \bar{M}(\xi(\omega), \theta^*)\, p(\theta^*)) \tag{6.52}$$

for some constant K which is independent of $\xi(\omega)$ and θ^*. It is seen that (6.52) has the form of the criterion considered in the previous Theorem. The difference is that now θ^* depends upon $\xi(\omega)$. But since (6.47) is still valid, we conclude that the set of all average information matrices is the convex hull of all average information matrices corresponding to single frequency designs, and the proof follows immediately. □

Since in (6.52) θ^* depends upon $\xi(\omega)$ this is a more complicated optimization problem than the previously considered. In a practical application, though, it is not necessary to actually find θ^*, one would use the result of the theorem and apply it to the criterion (6.51) directly. From the proof it is readily seen that in the general case (6.38) it is also possible to find an optimal design comprising a finite number of sinusoids, as long as ϕ is a continuous convex function of the information matrix.

The criteria discussed so far have all been based on ML estimation of the parameters. If instead a MAP estimator is used to estimate the unknown parameters the criteria for optimality must be changed accordingly. The following theorem establishes a relation between the posterior covariance of the parameters and Fisher's information matrix when a MAP estimator is used.

Theorem 6.3 *Assume that the prior knowledge about the model parameters is embodied in a Gaussian distribution with covariance matrix $\boldsymbol{\Sigma}_{\mathrm{pre}}$. Also assume that a MAP estimator is used to estimate the un-*

known parameters based on sampled observations for a model brought into the regression form

$$y_t = \varphi_t^T \theta + \epsilon_t$$

where $\{\epsilon_t\}$ *is a sequence of Gaussian random variables with known covariance, uncorrelated with* $\{\varphi_t\}$. *Then the posterior covariance matrix* $\mathbf{\Sigma}_{post}$ *is given by*

$$\mathbf{\Sigma}_{post}^{-1} = \mathbf{\Sigma}_{pre}^{-1} + \mathbf{M}_F = \mathbf{\Sigma}_{pre}^{-1} + N\bar{\mathbf{M}}_F \tag{6.53}$$

$\mathbf{M}_F$ *is Fisher's information matrix,* $\bar{\mathbf{M}}_F$ *the average information matrix, and* N *the length of experiment.*

PROOF: *see (Sadegh et al., 1994).* □

In the following, the concept of information is related to Lindley's measure of average information. In this way, we are able to formulate design criteria also based on MAP estimators. First some definitions are needed.

Definition 6.1 *The entropy of a random variable* X *having probability density function* $p(X)$ *is defined as*

$$H_x = -E_X[\log p(X)] \ . \tag{6.54}$$

Definition 6.2 *Lindley's measure of the average amount of information provided by an experiment* ξ *with data* y *and parameters* θ *is defined as*

$$J(\xi) = H_\theta - E_y[H_{\theta|y}] \ . \tag{6.55}$$

Now the relation between Fisher's information matrix and Lindley's measure of average information can be established via the following theorem (Sadegh et al., 1994).

Theorem 6.4 *With the same assumptions as in Theorem 6.3, maximizing Lindley's measure of the average amount of information, $J(\xi)$, is equivalent to solving the optimization problem*

$$\begin{aligned} &\min_{\xi} J \\ &J = -E_{\theta}[\log\det\{\boldsymbol{\Sigma}_{pre}^{-1} + \mathbf{M}_F\}] \end{aligned} \tag{6.56}$$

PROOF: *The estimation is regarded as a means by which further information about the system parameters is provided. Since MAP is the mode of the posterior distribution, the maximum amount of information with respect to Lindley's measure is obtained by MAP. Since the prior distribution of the parameters and the distribution of observations given θ is Gaussian, the posterior distribution of the parameters is also Gaussian. Denote the posterior mean and covariance by $\bar{\theta}$ and $\boldsymbol{\Sigma}_{post}$. From (6.53) we have*

$$\boldsymbol{\Sigma}_{post}^{-1} = \boldsymbol{\Sigma}_{pre}^{-1} + \mathbf{M}_F$$

From Definition 6.2

$$J(\xi) = -E_{\theta}\{\log p(\theta)\} + E_y E_{\theta|y}\{\log p(\theta|\mathbf{y})\} \tag{6.57}$$

$$\begin{aligned} &= -E_{\theta}\{\log p(\theta)\} + E_y E_{\theta|y}\{-\frac{n}{2}\log(2\pi)\} \\ &\quad -E_y E_{\theta|y}\{\frac{1}{2}\log\det\Sigma_{post} + \frac{1}{2}(\theta-\bar{\theta})^T\boldsymbol{\Sigma}_{post}^{-1}(\theta-\bar{\theta})\} \end{aligned} \tag{6.58}$$

As all the other terms obviously are constants, we only focus on the last two terms

$$E_y E_{\theta|y}\{\frac{1}{2}\log\det\boldsymbol{\Sigma}_{post}\}$$

$$= E_\theta E_{y|\theta}\{\frac{1}{2}\log \det \boldsymbol{\Sigma}_{post}\} \tag{6.59}$$

$$= E_\theta\{\frac{1}{2}\log \det \boldsymbol{\Sigma}_{post}\} \tag{6.60}$$

$$= -E_\theta\{\frac{1}{2}\log \det \boldsymbol{\Sigma}_{post}^{-1}\} \tag{6.61}$$

The other term can be written as

$$E_y E_{\theta|y}\{\frac{1}{2}(\theta-\bar{\theta})^T \boldsymbol{\Sigma}_{post}^{-1}(\theta-\bar{\theta})\}$$

$$= E_y E_{\theta|y}\{\frac{1}{2}\mathrm{trace}\boldsymbol{\Sigma}_{post}^{-1}(\theta-\bar{\theta})(\theta-\bar{\theta})^T\} \tag{6.62}$$

$$= E_y\{\frac{1}{2}\mathrm{trace}\boldsymbol{\Sigma}_{post}^{-1}E_{\theta|y}[(\theta-\bar{\theta})(\theta-\bar{\theta})^T]\} \tag{6.63}$$

$$= E_y\{\frac{n}{2}\} = \frac{n}{2} \tag{6.64}$$

where n *is the number of parameters. Now using (6.53) establishes the theorem.* □

An approximation of the mean value in (6.56) is optained by setting the parameters equal to their prior mean values. This approximation, which corresponds to a local design, simplifies the computations considerably. Thus depending on the estimation method, ML or MAP, and the choice of local or average criterion the following criteria would be of interest:

$$\begin{aligned} J_1 &= -[\log \det(\bar{M}_F)]_{\theta=E\{\theta\}} \\ J_2 &= -[\log \det(N\bar{M}_F + \Sigma_{pre}^{-1})]_{\theta=E\{\theta\}} \\ J_3 &= -E_\theta[\log \det(\bar{M}_F)] \\ J_4 &= -E_\theta[\log \det(N\bar{M}_F + \Sigma_{pre}^{-1})] \end{aligned} \tag{6.65}$$

These criteria demonstrate different levels of including partial prior information about parameter values. Optimization with respect to J_1 results in designs which are strongly dependent upon the prior information. The dependence is even more pronounced in J_2. It may therefore be wise to perform a sensitivity analysis, i.e. determine the sensitivity of the design to changes in the parameters, when using these criteria. An alternative is to choose an average criterion like J_3 and J_4. Table 6.1 summarizes the relation between the choice of estimators and the optimization criterion, expressed in both a local and average form. Applications with these criteria

	ML	MAP
Local design	J_1	J_2
Bayesian/average	J_3	J_4

Table 6.1. Summary of the optimality criteria

are found in e.g. (Melgaard et al., 1993; Sadegh et al., 1994).

6.4 Sampling Time and Presampling Filters

When estimating the parameters in a continuous time model from sampled data, the parameters will in general depend on the input to the system, the sampling instants and the presampling filter. The aims of designing optimal sampling instants and presampling filters is to avoid loss of information from the data due to sampling. Consider the following scalar continuous time system

$$y(s) = G(s)\,u(s) + \eta(s)\ , \qquad s = j\omega\ , \tag{6.66}$$

where u and y are the input and output, respectively, and η denotes colored measurement noise having spectral density $\psi(\omega)$ for all $\omega \in]-\infty, \infty[$.

Assume that the input signal spectrum is band limited to $[-\omega_h, \omega_h]$. The case of uniform sampling interval is considered. Suppose that the output is sampled faster than the Nyquist rate for ω_h, i.e.

$$\omega_s > 2\,\omega_h \tag{6.67}$$

where $\omega_s = 2\pi f_s$ and f_s is the sampling frequency. Such a sampler does not distort the part of the output spectrum arising from the input. The part of the output spectrum arising from the noise will, however, be distorted due to aliasing. The aliased noise spectrum is a superposition of different parts of the original spectrum

$$\psi_s(\omega) = \psi(\omega) + \sum_{r=1}^{\infty} (\psi(\omega + r\omega_s) + \psi(\omega - r\omega_s)) \tag{6.68}$$

for $\omega \in]-\omega_s/2, \omega_s/2[$. From expression (6.68) it is seen that $\psi_s(\omega) - \psi(\omega)$ is positive definite. It follows that the experiment with sampled data cannot be better than the corresponding experiment with continuous observations, see (Goodwin & Payne, 1977). However by inclusion of a suitable presampling filter, equality can be achieved.

Assume the presampling filter has the transfer function F with the ideal property

$$|F(j\omega)| = 0\,, \qquad \text{for } \omega \le -\omega_s/2 \text{ or } \omega \ge \omega_s/2\,, \tag{6.69}$$

and invertible otherwise. By including the presampling filter the filtered and sampled noise spectrum $\psi_{fs}(\omega)$ is

$$\psi_{fs}(\omega) = \psi_f(\omega) + \sum_{r=1}^{\infty} (\psi_f(\omega + r\omega_s) + \psi_f(\omega - r\omega_s)) \tag{6.70}$$

and $\psi_f(\omega)$ is the filtered noise spectrum given by

$$\psi_f(\omega) = F^T(j\omega)\psi(\omega)F(-j\omega)\,. \tag{6.71}$$

For a presampling filter satisfying (6.69), $\psi_{fs}(\omega)$ reduces to $\psi_f(\omega)$. This shows that the specified ideal presampling filter F eliminates the information loss due to sampling.

The approach discussed so far for optimal design of sampling time and presamplingfilter has considered the average information matrix per unit time. This is useful for design of experiments with a fixed experiment *time*, but without constraints on the number of samples. This leads to separate design of optimal input followed by choice of sampling time and presampling filter according to (6.67) and (6.69). However, if the total number of samples is constrained, then it is more appropriate to consider the average information matrix per *sample*. If the input spectrum is band limited to $[-\omega_h, \omega_h]$ and the sampling time and presampling filter is chosen according to (6.67) and (6.69) as before, then the average information matrix per sample is given by

$$\tilde{M}_F = \bar{M}_F / f_s \, , \tag{6.72}$$

where $\bar{M}_F$ is the average information matrix per unit time and f_s is the sampling rate. Optimizing a criterion based on this expression leads to a joint optimal design of input and sampling time for experiments with constrained number of samples. Such a design will in general lead to a compressed optimal input spectrum with a lower sampling rate and hence increased experiment time compared with the continuous observation case (Goodwin & Payne, 1977).

6.5 Example

In the following a simple example is given on the use of some of the different criteria for optimal design of input signal. The optimal design of sampling time is not covered by this example, a fixed sampling time is used. Consider a first order stochastic differential equation with discrete observations:

$$dx(t) = -ax(t)\,dt + bu(t)\,dt + dw(t)\ , \tag{6.73}$$

$$y_k = x_k + e_{2,k}\ , \tag{6.74}$$

where $w(t)$ is a Wiener process with $E(dw(t)) = 0$, $V(dw(t)^2) = r\,dt$, and $e_{2,k}$ is a Gaussian white noise process with $V(e_{2,k}) = r_2$. $dw(t)$ and $e_{2,k}$ are assumed to be independent; subscript k is shorthand for t_k. We are interested in the set of parameters $\theta = [a, b]'$, and to find an optimal input signal for the system. Since the inputs are assumed to be constant within a sampling interval, the corresponding discrete time model is written:

$$x_{k+1} = \alpha x_k + \beta u_k + e_{1,k}\ , \tag{6.75}$$

where α, β and $V(e_{1,k})$ depends on the sampling time τ by the following expressions:

$$\alpha = e^{-a\tau}\ , \qquad \beta = \int_0^\tau e^{-as} b\,ds = \frac{b}{a}(1 - e^{-a\tau})\ ,$$

$$V(e_{1,k}) = \int_0^\tau e^{-as} r e^{-as}\,ds = \frac{r}{2a}(1 - e^{-2a\tau})\ .$$

The discrete time model is brought into the innovations form

$$\hat{x}_{k+1} = \alpha\hat{x}_k + \beta u_k + K\epsilon_k\ , \tag{6.76}$$

$$y_k = \hat{x}_k + \epsilon_k\ , \tag{6.77}$$

where

$$V(\epsilon_k) = \bar{P} + r_2\ , \tag{6.78}$$

$$K = \alpha\bar{P}(\bar{P} + r_2)^{-1} , \tag{6.79}$$

and $\bar{P}$ is the positive semidefinite solution of the stationary Ricatti equation,

$$\bar{P} = \alpha^2\bar{P} + r_1 - (\alpha\bar{P})^2(\bar{P} + r_2)^{-1} . \tag{6.80}$$

The innovations form of the model (6.76) and (6.77) may be rewritten as:

$$y_k = \frac{\beta q^{-1}}{1 - \alpha q^{-1}} u_k + \frac{1 + (K - \alpha)q^{-1}}{1 - \alpha q^{-1}} \epsilon_k . \tag{6.81}$$

Since equation (6.81) is of the form (6.40) we are now ready to formulate the average Fishers information matrix for the problem. Considering a power constrained input in the frequency domain we use equations (6.47), (6.48) and (6.49).

By considering the criterion $\min_\xi J_1 = \min_\xi(-\log\det \bar{M}_F)$, Theorem 6.1 states that it suffices to look for designs comprising no more than 3 sinusoids for this example,

$$\bar{M}_F(\xi) = \sum_{k=1}^{3} \mu_k \bar{M}_F(\omega_k) , \tag{6.82}$$

where $\mu_k \geq 0$ and $\sum \mu_k = 1$. $\bar{M}_F(\omega_k)$ can be separated into $\tilde{M}_F(\omega_k)$ and $\bar{M}_c$ according to (6.48) and (6.49). It should be remarked, that when a local design is considered, only $\bar{M}_F(\omega_k)$ need to be calculated, but when a Bayesian design is used the total $\bar{M}_F(\omega_k)$ must be calculated. For the calculation of the matrices in this example the computer algebra language, MAPLE, was used. Hence the optimization problem is formulated and some general available software for solving nonlinearly constrained minimization can be applied.

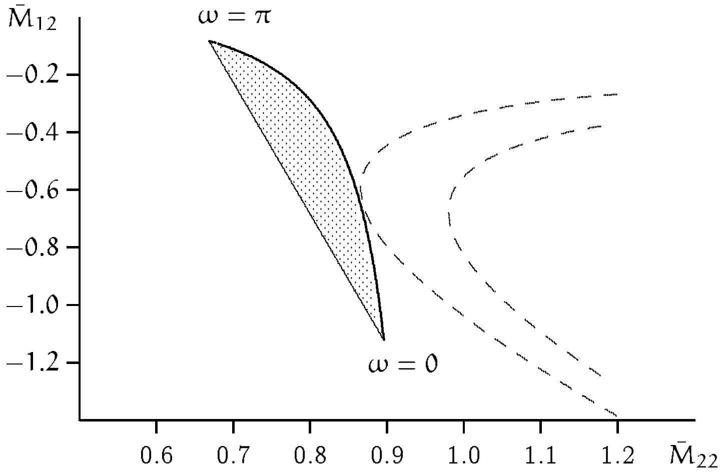


Figure 6.4. *Convex hull of single frequency designs for the example and contours of constant* det $\bar{M}_F$ *(increasing to the right).*

It turns out, that a single sinusoid is sufficient for optimality in the considered case. A plot of this optimal input frequency for different values of a is shown below for $b = 1.0$, $r = 1$, $r_2 = 0.1$ and $\tau = 1$.

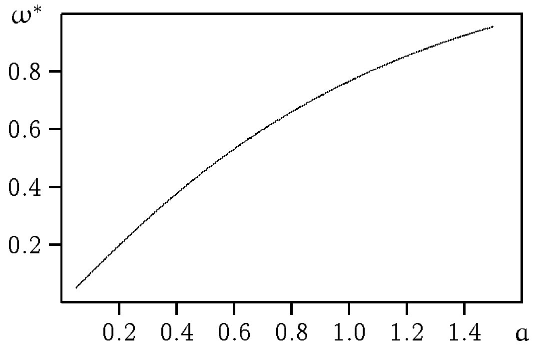


Figure 6.5. *Optimal design of input frequency* ω^* *for different values of* a.

Now we like to compare this traditional local design of experiments with the Bayesian approach for a given prior probability density of the parameters. Hence we consider the criterion $\min_\xi J_2 = \min_\xi E_\theta(-\log\det \bar{M}_F)$. It is assumed that parameter a is uniformly distributed in the interval $[0.5, 1.1]$ and the other parameters are fixed to the same values as above. The optimal design in this case is

$$\omega_2^* = 0.4575\,, \qquad J_2(\omega_2^*) = 1.1951\,. \tag{6.83}$$

This should be compared to the previous design using the expected value of the parameters, $E(a) = 0.8$, which has the optimal design, according to J_1

$$\omega_1^* = 0.6606\,, \qquad J_1(\omega_1^*) = 0.6107\,, \tag{6.84}$$

hence, here the two approaches give a different design. The average information when a Bayesian design is used is smaller than the information from the local design, when the used $E(a)$ is true.

Instead of optimizing the average performance with respect to the prior distribution one might be interested in a design which optimize the worst possible design. A plot of the performance of single frequency designs for different values of a is given below. Consider the minimax criterion

$$\min_\xi J_3 = \min_\xi(\max_{a\in\Theta} -\log\det \bar{M}_F) \tag{6.85}$$

Minimizing this criterion for the example with $a \in [0.5, 1.1]$ and the other parameters as above gives the optimal design

$$\omega_3^* = 0.8128\,, \qquad J_3(\omega_3^*) = 1.3741\,, \tag{6.86}$$

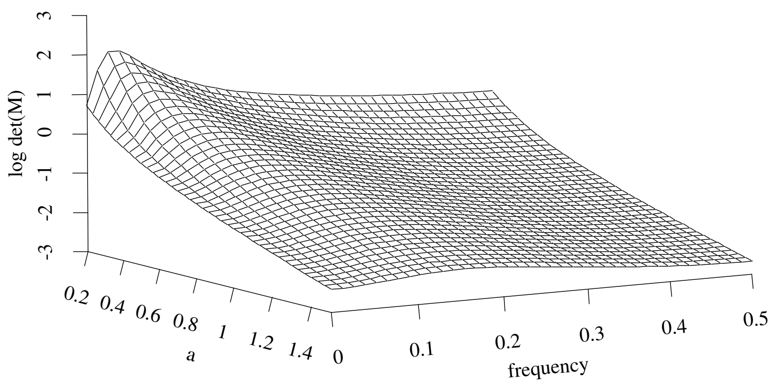


Figure 6.6. *The performance of single frequency designs for different values of* α. *The frequency axis has the scale* $f = 2\pi\omega$.

which is seen to differ from the previous designs. The value of the criterion expresses the worst possible performance from this design depending on the true value of α.

6.6 Summary

In this chapter the different aspects of experimental design has been discussed. Mainly the choice of optimal input signal is discussed. Different criteria for optimality of the design have been considered, mostly based on the information matrix as a characterization of the parameter uncertainty. A measure of information is selected as a function of the information matrix to represent the importance of the different parameters of the model or a transformation of these parameters. This measure can either be evaluated as a local design, i.e. at some fixed value of the parameters, e.g. the prior mean, or the criterion can be treated in a Bayesian approach by using the expectation of the measure with respect to the prior distribution of the parameters. Finally a minimax criterion can be defined for the measure of information. This last class of criteria is the least sensitive against single non-informative experiments.

Chapter 7

Case #1 Building Performance

The application considered in this chapter is related to the Commission of the European Communities' (CEC) research project called PASSYS. The aim of this project is to establish a common basis within the European Community for determining the energy dynamics of building components, especially components related to passive solar energy. Passive solar design has been recognized as an important potential for energy conservation. Many new components and systems have been developed in the late 1970s and 1980s. However, little is known about their actual thermal and solar dynamic characteristics. A further uncertainty is their unknown performance when the components are applied to buildings and exposed to variations in climate.

Within the PASSYS project, a test procedure for building components, using short-term performance data from a test cell, is developed and defined. The south wall and, for some test cells, the roof are removable. The test cell is calibrated using a highly insulated opaque south wall. Different south wall components could then be inserted in place of the calibration wall. The south wall, which is the actual passive solar system, is fixed in an insulated frame. Any kind of wall can be incorporated in this frame. The

Figure 7.1. *The PASSYS test cell. The south wall is not as shown here installed directly in a steel frame, but in an insulated frame in order to decrease the thermal bridges.*

test cell has a test room of 13.8 m^2 ground surface and 38 m^3 air volume with an adjoining service room to the north, accommodating measuring and air-conditioning equipment. The U-value of the envelope is less than 0.1 W/m^2K. A further description of the test cell is found in (Wouters & Vandaele, 1990).

7.1 The Heat Diffusion Equation

In order to understand the derivation of physical models for buildings the heat transfer across a uniform slab of building material will be investigated. The formalism of (Sonderegger, 1977) and (Lindfors, Christoffersson, Roberts, & Anderlind, 1992) is followed. Under some ideal situation, the temperature $v(x, t)$ at time t and position x in the slab, is described by the one-dimensional heat conduction equation

$$\frac{\partial v(x,t)}{\partial t} = \alpha \frac{\partial^2 v(x,t)}{\partial x^2} \tag{7.1}$$

where $\alpha = \frac{\kappa}{\rho\chi}$ $[m^2/s]$ is the thermal diffusivity, κ $[W/(Km)]$ is the thermal conductivity, ρ $[kg/m^3]$ is the density, and χ $[J/(kgK)]$ is the specific heat capacity. By this ideal situation, corner effects, thermal bridges, radiation from surfaces and heat transport, such as moisture effects has been neglected, as well as other non-linear phenomena. Equation (7.1) is a linear time-invariant partial differential equation. This kind of models are also known as distributed parameter models, and they are characterized by a state vector, which is infinite dimensional. There are in principal two ways to deal with this kind of models. One approach is to replace the space derivative by difference expressions, which will approximate the PDE by a finite set of ordinary differential equations. Then a *lumped* parameter model is obtained, which means that the distributed states are lumped into a finite set. The other approach is to keep the PDE description and only in the final numerical stage, introduce some approximations to make the computations feasible.

The type of models that is used for building modeling described in the following sections are based on the lumped parameter approach. It will be demonstrated in this section how simple low order lumped models will approximate the diffusion equation (7.1) quite well, in realistic environments, where also noise is present.

The complete solution of (7.1) is determined if initial conditions and the boundary conditions at each surface of the slab are given

$$\nu(x, 0) \tag{7.2}$$

$$\nu(x_1, t) \quad \text{or} \quad q(x_1, t) \tag{7.3}$$

$$\nu(x_2, t) \quad \text{or} \quad q(x_2, t) \tag{7.4}$$

where

$$q(x, t) = -\kappa \frac{\partial \nu(x, t)}{\partial x} \tag{7.5}$$

denotes the heat flux, and x_1 and x_2 are the positions at the respective surfaces. The problem is sketched in Fig. 7.2.

It is now assumed that the initial conditions (7.2) has lost their influence on the solution. That is, we assume that the slab has been exposed to its ordinary environment for a suitable long time, related to its thickness, d, and its thermal diffusivity, α

$$t \gg \frac{d^2}{\alpha} \tag{7.6}$$

where the term on the right hand side, loosely speaking, is called the time constant of the slab. This assumption is also a requirement on the system to be operating under stationary conditions, which allows for a frequency representation of the model.

The Laplace (or Fourier) transform of (7.1) is given by

$$\frac{\partial^2 \nu(x, s)}{\partial x^2} = \frac{s}{\alpha} \nu(x, s) \tag{7.7}$$

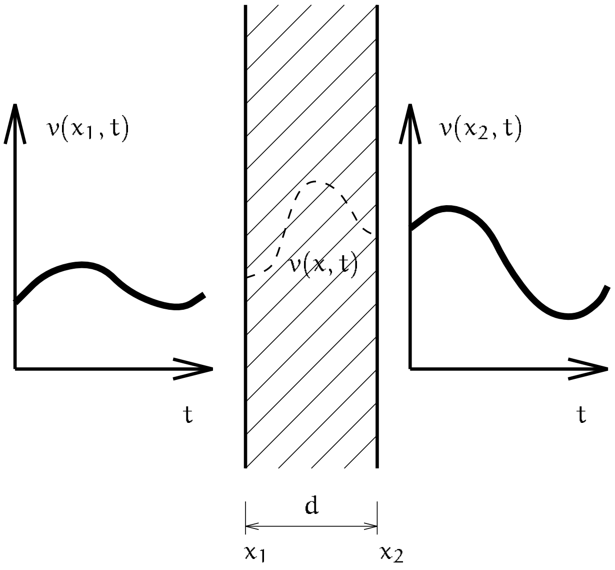


Figure 7.2. *Schematic temperature distribution and boundary conditions for the heat conduction through a wall.*

where s is the Laplace operator. By letting $s = j\omega$, where ω is the angular frequency and $j = \sqrt{-1}$, the Fourier transform is obtained. Equation 7.7 is an ordinary differential equation, which has the general solution

$$v(x, s) = F_1(s)\sinh(x\sqrt{s/\alpha}) + F_2(s)\cosh(x\sqrt{s/\alpha}) \tag{7.8}$$

where the functions $F_1(s)$ and $F_2(s)$ are found by matching (7.8) to the boundary conditions of (7.3) and (7.4). In principle Equation 7.8 provides the distribution of temperatures or heat fluxes across the entire wall, but usually one is only interested in the conditions on the boundaries, that is either surfaces of the wall. For that purpose it is possible to express the solution in the following matrix formulation, see (Sonderegger, 1977;

Lindfors et al., 1992)

$$\begin{bmatrix} v(x_1,s) \\ q(x_1,s) \end{bmatrix} = \begin{bmatrix} \cosh(d\sqrt{s/\alpha}) & \frac{\sinh(d\sqrt{s/\alpha})}{\kappa\sqrt{s/\alpha}} \\ \kappa\sqrt{s/\alpha}\sinh(d\sqrt{s/\alpha}) & \cosh(d\sqrt{s/\alpha}) \end{bmatrix} \begin{bmatrix} v(x_2,s) \\ q(x_2,s) \end{bmatrix} \tag{7.9}$$

where $q(x,s)$ is the Laplace transform of the heat flux. Normally, and in this case, we are not interested in the heat flux on the outside or cold side of the wall. If we denote x_1 as the inside and x_2 as the outside, Equation (7.9) can be reduced to a one dimensional equation, without $q(x_2,s)$. The relations can also be formulated in the thermal parameters we are interested in, namely

$$R = d/\kappa \tag{7.10}$$

$$C = d\rho\chi \tag{7.11}$$

where R is the thermal resistance $[Km^2/W]$ of the wall, the inverse of the U-value. C is the overall thermal capacity per unit area $[J/(Km^2)]$ of the wall. By using these parameters and reducing Equation (7.9) into one equation, we get

$$v_1(s) = \frac{\frac{R}{\sqrt{sRC}}\sinh(\sqrt{sRC})}{\cosh(\sqrt{sRC})}q_1(s) + \frac{1}{\cosh(\sqrt{sRC})}v_2(s) \tag{7.12}$$

which can also be written

$$v_1(s) = G_1(s)q_1(s) + G_2(s)v_2(s) \tag{7.13}$$

where $v_1(s)$ denotes $v(x_1,s)$ etc., and $G_1(s)$ is the transfer function between the heat flux and the indoor temperature and $G_2(s)$ is the transfer function between the outdoor temperature and the indoor temperature. Note that the transfer functions are analytic functions of s but not rational.

In the limit, where $|sRC| \to 0$, which can be thought of as for very thin, light and conductive walls or slow time variability of the boundary conditions, making $|\omega|$, and hence s small, Equation 7.12 becomes

$$q_1(s) = \frac{v_1(s) - v_2(s)}{R} \tag{7.14}$$

which is the well known static heat balance equation.

7.1.1 Lumped Parameter Models

For the purpose of simulating (or predicting) the dynamic behavior of a particular wall, one may wish to estimate the parameters in (7.12) from measurements of the heat flux and surface temperatures. This may not seem necessary for a homogeneous slab of well known material and dimension. But when the ideas are extended to cover multilayer walls and more complex systems, as a whole building, there are no other way than to estimate the parameters of the model from measurements, if an accurate model is needed. In the following models for the homogeneous slab are discussed, but it is kept in mind that the ideas will later be extended to more complex systems.

One possibility for estimating the parameters directly in (7.12) from measurements is to keep the calculations in the frequency domain, by Fourier transforming the time series of measurements. This approach has been used by (Lindfors et al., 1992) to find least-squares estimates of R and C for a homogeneous slab.

Another approach, which will be discussed here, is to approximate the partial differential equation by a set of ordinary differential equations and then estimate the parameters in the time domain. Usually a low order approximation is sufficient. The parameters of the original partial differential equation can easily be calculated from the parameters of the ordinary differential equations afterwards if that is of interest. The advantage of this approach is that it is easy to extend it to more complex systems and even consider non-linearities in the model. The resulting lumped parameter models are also often called R-C network models because they can be constructed from networks of thermal resistors and capacitors equivalent to an electrical network. Following the discussions in (Sonderegger, 1977) it shall be demonstrated how the exact transfer functions of (7.12) can be approximated by an equivalent R-C network.

A distributed system, as opposed to lumped-element systems, is characterized by an infinite number of poles and zeros of its transfer functions. The dynamic response of the system is determined by its poles and zeros, and therefore it is important how the approximating lumped model has its poles and zeros. It has been demonstrated by (Goodson, 1970) that by using the infinite product expansion technique the exact poles and zeros of the transfer functions are preserved for any order approximation. This imply that the criterion for determining the number of products to use (for deterministic models) is the frequency bandwidth of the desired model. For a stochastic system, where noise is present, there is an upper limit for the number of products, which is necessary for representing the system.

The transfer functions in (7.12) can be written as

$$v_1(s) = \frac{B(s)}{D(s)} q_1(s) + \frac{1}{D(s)} v_2(s) \tag{7.15}$$

with the nominator, $B(s) = \frac{R}{\sqrt{sRC}} \sinh(\sqrt{sRC})$ and the denominator, $D(s) = \cosh(\sqrt{sRC})$. Their zeros are put in evidence by their infinite product expansion:

$$B(s) = \frac{R}{\sqrt{sRC}} \sinh(\sqrt{sRC}) = R \prod_{n=1}^{\infty} \left(1 + \frac{sRC}{n^2\pi^2}\right) \tag{7.16}$$

$$D(s) = \cosh(\sqrt{sRC}) = \prod_{n=1}^{\infty} \left(1 + \frac{sRC}{(n - 1/2)^2\pi^2}\right) \tag{7.17}$$

which can be found in (Goodson, 1970).

We now wish to find the $R - C$ network that corresponds to a given truncation of Equation 7.16 and 7.17. The considered network models are sketched in Figure 7.3. A first order approximation corresponds to the solid components in the figure, the second order approximation is obtained by also adding the dashed components. The polynomials corresponding to

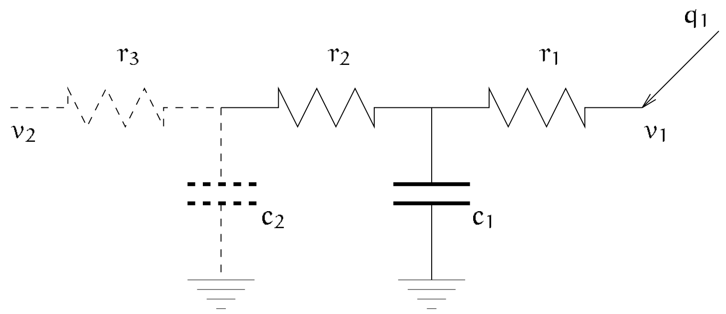


Figure 7.3. $R - C$ *network models of the wall. First (solid) and second (solid + dashed) order models are shown.*

(7.15) for the first order model is given by:

$$B_1(s) = (r_1 r_2 c_1)s + (r_1 + r_2) \tag{7.18}$$

$$D_1(s) = (r_2 c_1)s + 1 \tag{7.19}$$

Both polynomials are of first order, thus only the first zero from the infinite product expansion of each polynomial $B(s)$ and $D(s)$ can be matched, in addition to the gain of $B(s)$, $|B(s = 0)| = R$. The resistances and capacitance of the lumped model that preserves the gain and the first zeros of $B(s)$ and $D(s)$ are given by:

$$\text{first order model from infinite product expansion} \quad \left\{ \begin{array}{lcl} r_1 & = & 0.2500\,R \\ r_2 & = & 0.7500\,R \\ c_1 & = & 0.5404\,C \end{array} \right. \tag{7.20}$$

The polynomials corresponding to (7.15) for the second order model is given by:

$$\begin{aligned} B_2(s) &= (r_1 r_2 r_3 c_1 c_2)s^2 + (r_1 r_3 c_2 + r_1 r_3 c_1 + r_1 r_2 c_1 + r_2 r_3 c_2)s \\ &\quad + (r_1 + r_2 + r_3) \end{aligned} \tag{7.21}$$

$$D_2(s) = (r_2 r_3 c_1 c_2)s^2 + (r_3 c_2 + r_3 c_1 + r_2 c_1)s + 1 \tag{7.22}$$

Since both polynomials are of second order, we are able to match the two first zeros from the infinite product expansion of each polynomial $B(s)$ and $D(s)$, in addition to the gain of $B(s)$. The resulting resistances and capacitances of the second order lumped model that preserves the gain and the two first zeros of $B(s)$ and $D(s)$ are given by:

$$\text{second order model from infinite product expansion} \quad \left\{ \begin{array}{lcl} r_1 & = & 0.1406\,R \\ r_2 & = & 0.3125\,R \\ r_3 & = & 0.5469\,R \\ c_1 & = & 0.2882\,C \\ c_2 & = & 0.3705\,C \end{array} \right. \tag{7.23}$$

By using this technique for approximating the distributed model, we obtain the characteristics that $\sum_i r_i = R$, which also mean that the gain (dc-component) is preserved by the model. Furthermore we have that for the first order model $\sum_i c_i = 0.54\,C$, and for the second order model $\sum_i c_i = 0.66\,C$ (it can be shown that if the model order $\to \infty$ then $\sum_i c_i \to C$). The technique also imply that the criterion for determining the model order is given by the required bandwidth of the model.

The frequency response of the transfer functions $B(s)/D(s)$ and $1/D(s)$ for the exact solution as well as first and second order approximations using infinite product expansion are shown on Figure 7.4 and 7.5.

Another way to view a lumped parameter model is to consider directly the approximation of the space-derivative of the partial differential equation. Let us consider the general approximation

$$\left.\frac{\partial^2 v(x,t)}{\partial x^2}\right|_{x_2} \simeq \left(\frac{v(x_1,t)-v(x_2,t)}{\delta_1} - \frac{v(x_2,t)-v(x_3,t)}{\delta_2}\right)/\epsilon_1 \quad (7.24)$$

where $x_2 \in]x_1, x_3[$, $\delta_1 = x_1 - x_2$ and $\delta_2 = x_2 - x_3$. It is also assumed that $\varepsilon_1 \in]\min(\delta_1,\delta_2), \max(\delta_1,\delta_2)[$. By inserting this approximation in (7.1) we obtain

$$\frac{\partial v(x_2,t)}{\partial t} \simeq$$
$$\frac{d^2}{RC}\frac{1}{\varepsilon_1}\left(\frac{1}{\delta_1}v(x_1,t) - (\frac{1}{\delta_1}+\frac{1}{\delta_2})v(x_2,t) + \frac{1}{\delta_2}v(x_3,t)\right) \quad (7.25)$$

since $\alpha = d^2/(RC)$. By comparing Equation 7.25 with the similar expression for the $R-C$ network as Fig. 7.3 it is seen that they match with the following terms:

$$r_i = R\frac{\delta_i}{d} \quad (7.26)$$
$$c_i = C\frac{\varepsilon_i}{d} \quad (7.27)$$

hence r_i can be considered as a fraction of the total R, directly proportional to a distance of the material, and since $\sum_i \delta_i = d$ we have that $\sum_i r_i = R$. The same holds for c_i except that $\sum_i \varepsilon_i < d$, and only in the limit we have $\sum_i c_i \to C$.

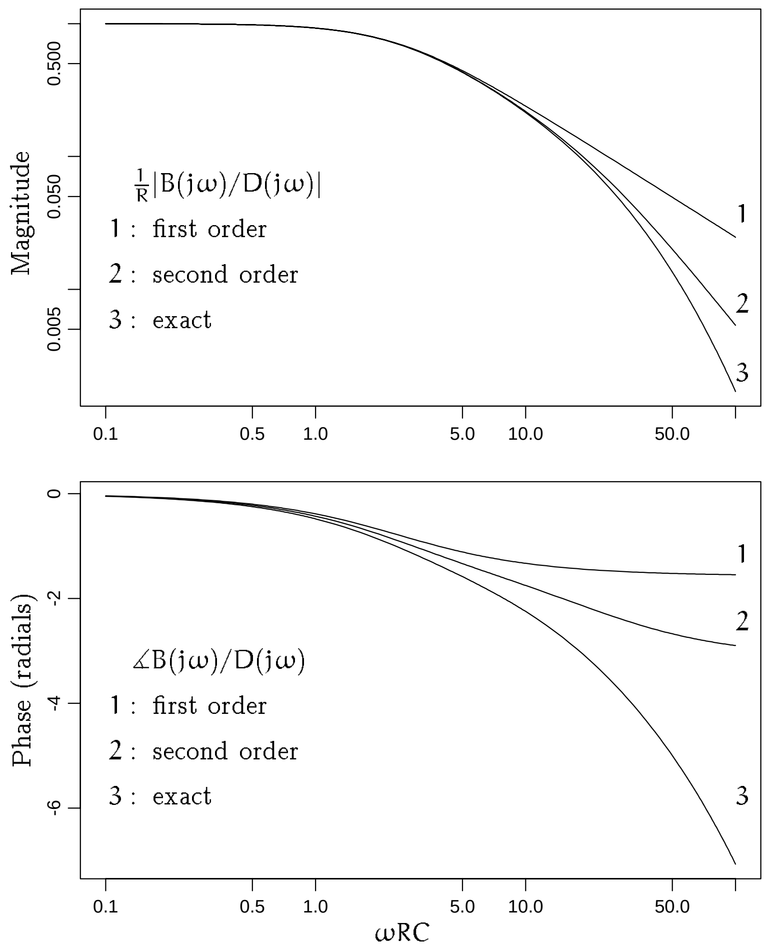


Figure 7.4. *Frequency response of* $B(j\omega)/D(j\omega)$ *shown for the first and second order* $R - C$ *approximation as well as the exact model.*

7.1.2 Measured Data

Attention is now focused to the model formulation (7.13), but with the extension that all the signals have been measured through some measuring

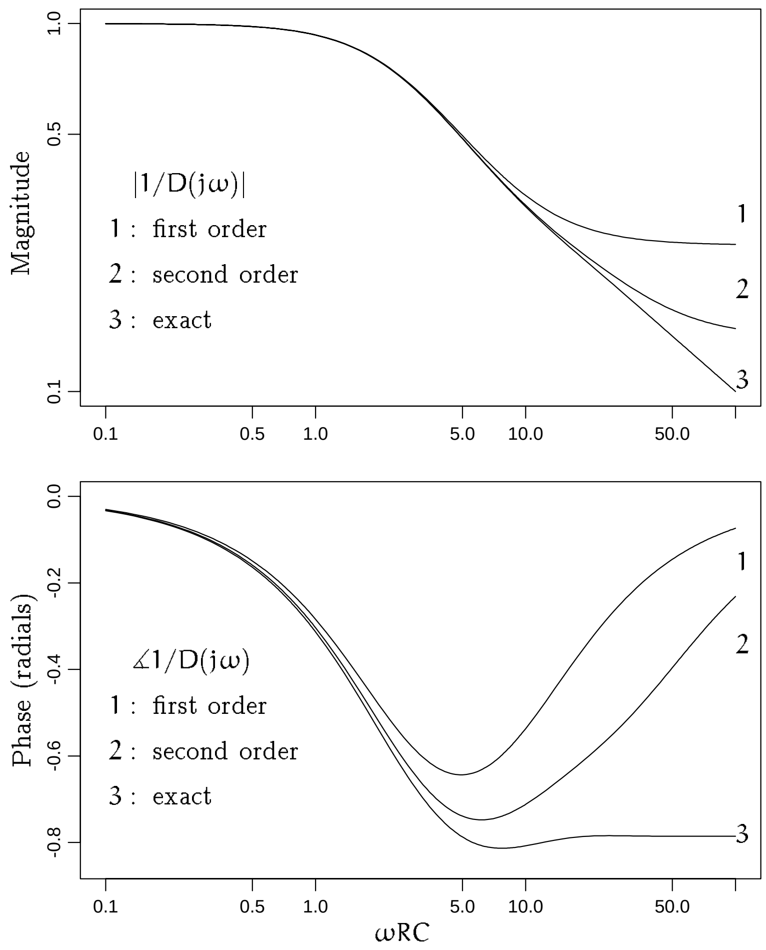


Figure 7.5. *Frequency response of* $1/D(j\omega)$ *shown for the first and second order* $R - C$ *approximation as well as the exact model.*

device. This means that all signals are corrupted by noise. For simplicity we consider all the signals to be disturbed by additive white noise, i.e.

$$v_1(s) = v_1'(s) + \epsilon_{v1}(s) \tag{7.28}$$

$$v_2(s) = v_2'(s) + \epsilon_{v2}(s) \tag{7.29}$$

$$q_1(s) = q_1'(s) + \epsilon_{q1}(s) \tag{7.30}$$

where $v_1'(s)$ denotes the true (undisturbed) signal and $\epsilon_{v1}(s)$ is the measurement noise, assumed to be zero mean Gaussian i.i.d. with standard deviation σ_{v1}, etc. For simplicity, the calculations in this section is kept in the continuous time domain, but the results also apply to discrete time if the data are sampled properly. This imply that the information loss, due to the sampling, is minimized by selecting a proper sampling time and anti-aliasing filter, see Section 6.4. By inserting (7.28) - (7.30) in (7.13) we obtain the model formulation:

$$v_1'(s) = G_1(s)q_1'(s) + G_2(s)v_2'(s) + N(s) \tag{7.31}$$

where the noise term is

$$N(s) = -\epsilon_{v1}(s) + G_1(s)\epsilon_{q1}(s) + G_2(s)\epsilon_{v2}(s) \tag{7.32}$$

A statistical tool which is useful for analyzing the relationship between a variable and some other variables is the *multiple coherency*, see Section 5.2.3. In this case the multiple coherency, $W^2_{v_1 q_1 v_2}(\omega)$, between $v_1(t)$ and $q_1(t)$, $v_2(t)$ represents the proportion of the power of $v_1(t)$ which can be explained by the relationship with $q_1(t)$ and $v_2(t)$, at frequency ω. It is given by the relation:

$$\Gamma_N(\omega) = \Gamma_{v_1}(\omega)[1 - W^2_{v_1 q_1 v_2}(\omega)] \tag{7.33}$$

Of course the shape of $W^2_{v_1 q_1 v_2}(\omega)$, and hence the ability to identify the transfer functions, is dependent upon the actual experiment, the sensors etc. Usually though the sensors will have the frequency characteristics of a low pass filter.

Example 7.1 *In this example we consider the situation where the* RC *of the component (wall) is large compared to the bandwidth of the input signals to the system, i.e.* $q_1(t)$ *and* $v_2(t)$. *This situation could be the case for a heavy, well insulated wall exposed to outdoor climate,* $v_2(t)$, *and a fast responding indoor heat supply. The frequency distribution is assumed to be uniform, in the considered scale:*

$$q_1(\omega) = 1 \tag{7.34}$$

$$v_2(\omega) = 1 \tag{7.35}$$

Furthermore, 4 different levels of white noise added to all the signals have been considered, cf. (7.28) - (7.30). The following 4 levels of the noise have been analyzed:

$$1: \quad \sigma^2_{v1}(\omega) = \sigma^2_{v2}(\omega) = \sigma^2_{q1}(\omega) = 0.0001 \tag{7.36}$$

$$2: \quad \sigma^2_{v1}(\omega) = \sigma^2_{v2}(\omega) = \sigma^2_{q1}(\omega) = 0.001 \tag{7.37}$$

$$3: \quad \sigma^2_{v1}(\omega) = \sigma^2_{v2}(\omega) = \sigma^2_{q1}(\omega) = 0.01 \tag{7.38}$$

$$4: \quad \sigma^2_{v1}(\omega) = \sigma^2_{v2}(\omega) = \sigma^2_{q1}(\omega) = 0.1 \tag{7.39}$$

In Figure 7.6 the squared coherency is shown for the given example, with increasing level of the noise. For a low level of the noise (curve no. 1*) the squared coherency is almost equal to* 1 *at all frequencies, in the considered scale. This is expected, since if there where no noise at all, then from Equation 7.33 we have that* $W^2_{v_1 q_1 v_2}(\omega) = 1$ *for all frequencies, where* $\Gamma_{v_1}(\omega) \neq 0$. *As the level of the noise is increased, the correlation between the inputs and the output decreases at high frequencies. This can be explained by the low-pass nature of the system, see Figure 7.4 and 7.5. The output* v_1 *is dominated by low frequen-*

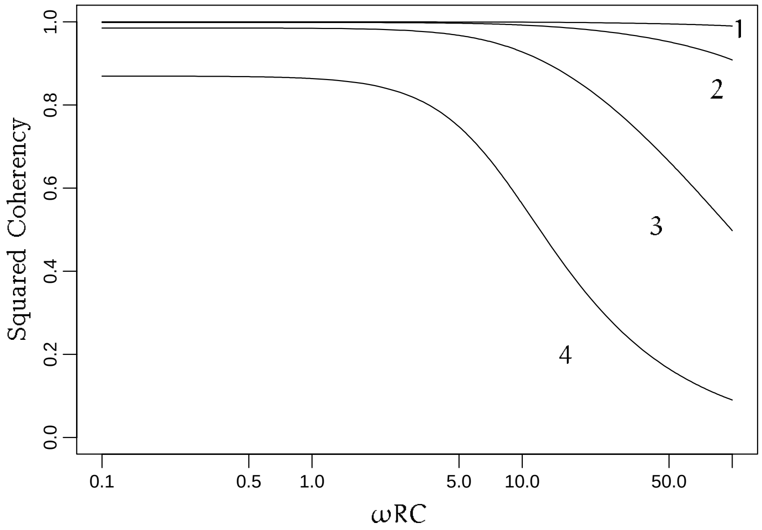


Figure 7.6. Squared coherency.

cies, whereas the noise on the output, ϵ_{v1} *is white, i.e. represented uniformly at all frequencies. Therefore we have the high values of the squared coherency spectrum at low frequencies.*

In the first example the inputs were supposed to be represented uniformly at all frequencies, relative to the RC of the system. For in-situ applications of building components it is more likely that the power spectrum of the inputs have a low frequency dominance relative to the system, and a peak at the frequency corresponding to the daily cycle of the climate. This situation is considered in the next example.

Example 7.2 *In this example the power spectra of the inputs are assumed to have a low frequency dominance, relative to the system dynamics. This could be the case for a light-weight construction with thin* *insulation. In the considered scale of frequencies, the power spectra of the input signals are:*

$$q_1(\omega) = 1/\omega^2 \tag{7.40}$$

$$v_2(\omega) = 1/\omega^2 \tag{7.41}$$

The different levels of additive white noise are the same as in the previous example. As can be seen from Figure 7.7 the coherency almost disappears at high frequencies and this is especially true for large noise levels (curve no. 4*).*

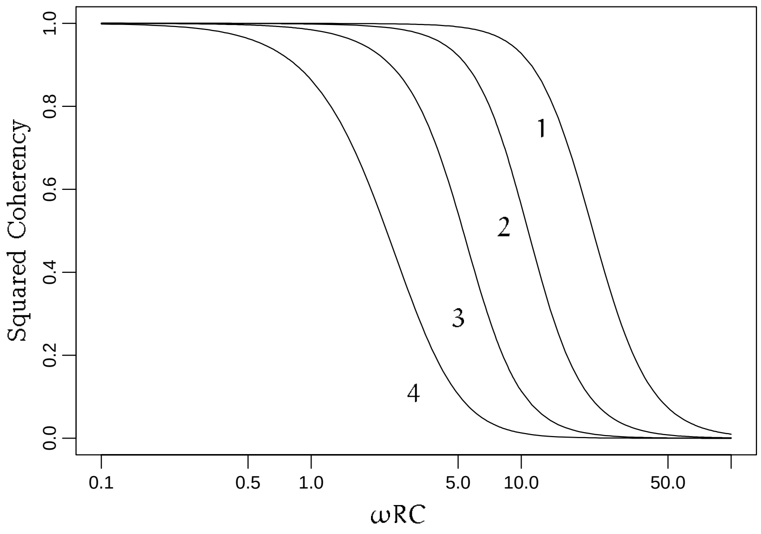


Figure 7.7. Squared coherency .

In both of these examples the data series are supposed to be measured. For this reason, measurement noise was taken into account. In real application the data series will also be of *finite length.* This imply that the squared coherency determined from the data, has some associated uncertainty. It is then possible to test for zero coherency (at frequency ω). From Priestley (1981) pp. 706 we have that, under the hypothesis that $W^2_{v_1q_1v_2}(\omega) = 0$,

$$\frac{2m\widehat{W}^2_{v_1q_1v_2}(\omega)}{(1 - \widehat{W}^2_{v_1q_1v_2}(\omega))} = F_{2,4m} \tag{7.42}$$

where m is related to the spectral estimate, using a Daniell window. This spectral estimate is computed by averaging $(2m+1)$ periodogram ordinates. For a fixed length of data, N, the parameter m must be chosen as a trade off between the bias and variance of the spectral estimate. As $m \uparrow$, the variance of the estimate $\downarrow$, and the bias $\uparrow$, and vice versa. If the shape of the true spectrum is known, then it is possible to design more optimal parameters for the window smoother, see (Priestley, 1981). The main point in this discussion is that for finite length data series there is a limit > 0, determined from Equation 7.42, below which we can not reject that the coherency is equal to zero at the given frequencies. Because the curves in Figure 7.6 and 7.7 are monotonous decreasing, such a limit would determine a certain frequency above which we could not reject that the coherency is equal to zero. This further imply, that if a model is fitted to the data, we are not able to distinguish between different models above this upper frequency limit. By comparing the frequency responses of the models in Figure 7.4 and 7.5 we conclude that it is not possible to distinguish between the exact model and a finite order lumped parameter model, if the models are to be fit to measured data of finite length.

The introduction of the infinite product expansion as an approximation to the exact model was merely done to argue for the structure of the $R-C$ network models in Figure 7.3, and demonstrate the capability of the $R-C$ network models to fit the exact model within a required bandwidth. In a real application one would prefer to fit the individual parameters of the $R-$

C network instead of estimating R and C directly from the infinite product expansion (7.20) or (7.23) or higher order. There are more reasons for this statement. Maybe some of the assumptions and simplifications used for the derivation of the exact model do not fully hold, e.g. the assumption of only one dimensional heat losses. The infinite product expansion approximates the exact model best at low frequencies and then it is worse as the frequency increases. This is usually what is needed, compare with the examples in this section, but normally the spectral distribution of the measured data is not as smooth as in the examples, e.g. the climate data have a peak around the frequency for the daily cycle. By estimating the individual parameters some extra degrees of freedom allows for a better fit to the data, which is necessary for the above reasons.

7.2 Validation of the Estimation Tool

The thermal characteristic of buildings is frequently approximated by a simple network with resistors and capacitances, see for instance (Sonderegger, 1977; Subbarao, 1985; Hammersten, van Hattem, Bloem, & Colombo, 1988; Madsen & Holst, 1993). In this section, such a (simplified) lumped parameter model for the dynamics of the test cell is presented. The model is used for simulations and estimations in the following sections. The objective of this work is to validate the tool used for estimation. If a number of stochastic independent data sets are simulated from a known model, and later the parameters of this model is estimated by the given tool, it is possible to validate the estimated parameters and the estimated uncertainty of the parameters. This will require from the estimation method that it is able to handle the considered type of model, but also that the method is implemented correctly on the computer, taking properly care of the numerics.

The dominating heat capacity of the test cell is located in the outer wall. For such buildings, the model with two time constants shown in Figure 7.8 is frequently found adequate. The states of the model are given by

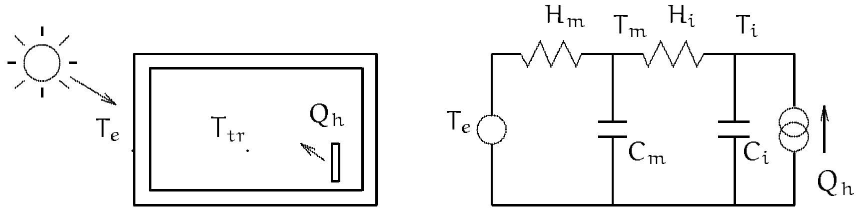


Figure 7.8. *A model with two time constants of the test building and the equivalent electrical network.*

the temperature, T_i, of the indoor air and possibly inner part of the walls with heat capacity C_i, and by the temperature, T_m, of the heat accumulating medium, with the heat capacity C_m. H_i is the transmittance of the heat transfer between the room air and the walls, while H_m is the heat transmittance between the inner part of the walls and the external surface of the walls. The input to the system is the heat supply, Q_h, and the outdoor surface temperature, T_e. The test cell is covered outside by a thin metal shell on which a large number of thermocouples are mounted (5 on each surface). T_e is calculated as an area-weighted mean of all the sensor signals. By considering the outdoor *surface* temperature instead of the outdoor *air* temperature as the input, the effect of solar radiation is automatically taken into account. The assumption of no feedback from the test cell to the external temperature is still valid because there is always some wind around the test cell, which will assure that the external surface temperature is determined from the external air temperature and the solar radiation on the surface alone.

In state space form the model is written,

$$\begin{bmatrix} dT_i \\ dT_m \end{bmatrix} = \begin{bmatrix} -H_i/C_i & H_i/C_i \\ H_i/C_m & -(H_i + H_m)/C_m \end{bmatrix} \begin{bmatrix} T_i \\ T_m \end{bmatrix} dt + \begin{bmatrix} 0 & 1/C_i \\ H_m/C_m & 0 \end{bmatrix} \begin{bmatrix} T_e \\ Q_h \end{bmatrix} dt + \begin{bmatrix} dw_i(t) \\ dw_m(t) \end{bmatrix} . \tag{7.43}$$

An additive noise term is introduced to describe deviations between the model and the true system, the term can also be considered as noise on the input signals. Hence, the model of the heat dynamics is given by the (matrix) stochastic differential equation

$$d\mathbf{T} = \mathbf{A}\mathbf{T}\,dt + \mathbf{B}\mathbf{U}dt + d\mathbf{w}(t) , \tag{7.44}$$

where $w(t)$ is assumed to be a Wiener process with incremental covariance matrix

$$\Sigma = \begin{bmatrix} \sigma_{1,i}^2 & 0 \\ 0 & \sigma_{1,m}^2 \end{bmatrix} . \tag{7.45}$$

The measured air temperature is naturally encumbered with some measurement errors, and hence the measurement equation is written

$$T_{tr}(t) = [1\ 0] \begin{bmatrix} T_i \\ T_m \end{bmatrix} + e(t) , \tag{7.46}$$

where $e(t)$ is the measurement error, assumed to be normally distributed with zero mean and variance σ_2^2.

The following values of the parameters have been estimated in an earlier experiment on a test cell: $H_i = 55.29$ W/K, $H_m = 13.86$ W/K, $C_i = 325.0$ Wh/K, $C_m = 387.8$ Wh/K, $\sigma_{1,i}^2 = 0.00167\ \mathrm{K}^2$, $\sigma_{1,m}^2 = 0.00978\ \mathrm{K}^2$, and $\sigma_2^2 = 0.00019\ \mathrm{K}^2$. Corresponding to these parameters, the time constants of the system are $\tau_1 = 3.03$ hours and $\tau_2 = 54.28$ hours.

Using these parameters for the model, we are able to simulate the sample paths of the system. By simulating several sample paths from the system and estimating the parameters of the model from each sample path this is a way of validating the estimation procedure. By considering several simulated sequences this investigation considers both the mean values and the variances of the estimated parameters. For the simulation study the following input and output signals of the model have been used: T_e is measured surface temperature, from a Danish test building, Q_h the heat supply, is a PRBS (pseudo-random binary sequence) with the number of stages, $n = 6$ and smallest switching interval $T_{prbs} = 8$ hours, (Godfrey, 1980), switching between 0 W and 300 W, and T_{tr} is the indoor room temperature, simulated from the specified model. In this study the sampling

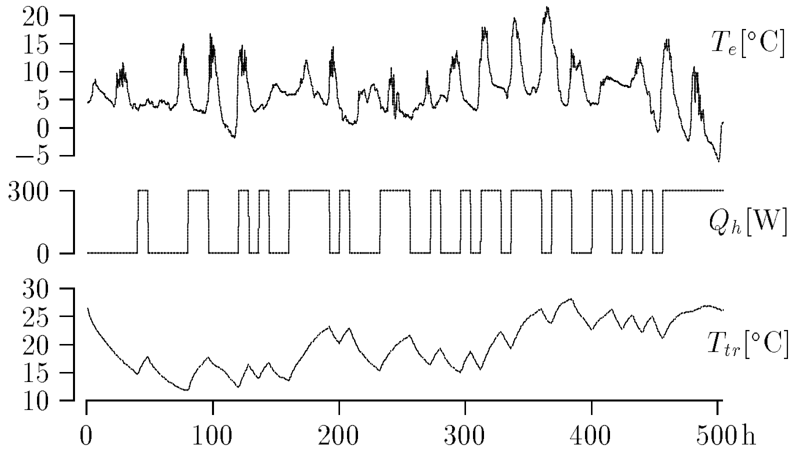


Figure 7.9. *The signals in the simulations: the outdoor surface temperature, the controlled heat power and the indoor air temperature.*

time is fixed to $T_{sampl} = 20$ minutes, and the length of each experiment is 21 days, which is equal to 1512 observations per simulated series. We have simulated 50 equal series, but with different realizations of the noise sequences. Refer to Section 2.5 for the details about simulating stochastic differential equations and noise generation.

In Table 7.1 and Table 7.2 the results are summarized from estimations using CTLSM (Melgaard & Madsen, 1993). This is a tool for maximum likelihood estimation of multivariate stochastic differential equations with a linear or non-linear state space formulation. In the tables results are summarized for the parameters of the model and for two physical characteristics, which can be calculated as functions of the model parameters. The UA-value is the overall heat transmittance and CI is the internal heat capacity of the building. For all parameters, the mean of the estimated values are given, which can be compared to the simulated values. Also the empirical variance and the mean of the estimated variances of the parameters are given. A comparison of these values will indicate if the method is able to estimate the right uncertainty of the parameters.

In Table 7.1 the results are obtained using 1-step predictions in the criterion for maximum likelihood. When estimating the true model, which is the case here, using 1-step predictions is the optimal choice (Kabaila, 1981; Stoica & Nehorai, 1989). On the other hand if the true model is not contained in the model set, it can be advantageous to filter the residuals before estimating the parameters. One way of doing this is to use a k-step prediction-horizon in the criterion for some $k > 1$. This kind of filtering works like a low-pass filter and will put more weight on the low frequency part of the dynamics. In the considered case the physical characteristic parameters UA- and CI-values represent low frequency dynamics. The columns of the tables are: x_{sim}, the simulated values, $\bar{x}$, the mean of the estimated values, s_x^2 is the empirical variance of the estimated parameters,

Parameter	x_{sim}	$\bar{x}$	s_x^2	$\bar{s^2}$	F-stat.	$\lvert t\rvert$-stat.
H_i [W/K]	55.290	55.279	2.1645	2.3772	0.9105	0.053
H_m [W/K]	13.860	13.867	0.0360	0.0390	0.9244	0.261
C_i [Wh/K]	325.00	325.0	9.5667	10.746	0.8902	0.011
C_m [Wh/K]	387.78	387.33	162.26	159.89	1.0148	0.250
$\sigma_{1,i}^2$ [K^2]	1.670e-3	1.650e-3	2.185e-8	2.475e-8	0.8825	0.975
$\sigma_{1,m}^2$ [K^2]	9.780e-3	9.554e-3	8.650e-7	1.109e-6	0.7803	1.716
σ_2^2 [K^2]	1.900e-4	1.912e-4	7.956e-10	7.446e-10	1.0686	0.308
UA [W/K]	11.082	11.084	0.0112	0.0111	1.0114	0.125
CI [MJ/K]	2.2862	2.2846	1.4553e-3	1.3815e-3	1.0535	0.304

Table 7.1. *Results from estimation of $n_e = 50$ series, using 1-step predictions in the criterion. The mean autocorrelation of residuals is,* $\bar{\rho}(1) = -0.001$.

$\bar{s^2}$ is the mean of the estimated variance of the parameters, F-stat. is an F statistic given by $Z_F = s_x^2/\bar{s^2}$ and $|t|$-stat. is a t statistic given by $Z_t = |\bar{x} - x_{sim}|/(s_x\sqrt{n_e})$.

In order to verify if the variance of the parameters provided by the estimation tool is equal to the empirical variance, one wish to test the hypotheses

$$\begin{aligned} H_0 : &\quad s_x^2 = \bar{s^2} \\ H_1 : &\quad s_x^2 \neq \bar{s^2} . \end{aligned}$$

Under H_0 we have in this case that $Z_F \sim F(49, \infty)$. The critical set for this test is $\{z < F(49,\infty)_{\alpha/2} \vee z > F(49,\infty)_{1-\alpha/2}\}$ on level α. By choosing $\alpha = 0.1$ we obtain the critical set $\{z < 0.74 \vee z > 1.35\}$. It is seen from Table 7.1 that we cannot reject H_0 for any parameter on the chosen level.

Another test is performed in order to verify that the estimated parameters are not biased. The following hypotheses are tested

$$H_0: \quad \bar{x} = x_{sim}$$
$$H_1: \quad \bar{x} \neq x_{sim}\,.$$

Under H_0 the distribution of the test statistic is $Z_t \sim t(49)$. The critical set for this test is $\{z > t(49)_{1-\alpha/2}\}$ on level α. For $\alpha = 0.1$, the critical set is $\{z > 2.0\}$, thus from Table 7.1 we cannot reject H_0 for any of the parameters on the chosen level.

Parameter	x_{sim}	$\bar{x}$	$s_{\bar{x}}^2$	$\bar{s^2}$	F-stat.	$\lvert t\rvert$-stat.
H_i [W/K]	55.290	55.339	2.2259	0.6154	3.617	0.232
H_m [W/K]	13.860	13.863	0.0383	0.0099	3.877	0.109
C_i [Wh/K]	325.00	324.8	9.486	3.3338	2.842	0.481
C_m [Wh/K]	387.78	387.20	163.74	47.685	3.4338	0.3210
$\sigma_{1,i}^2$ [K^2]	1.670e-3	1.616e-3	4.158e-8	1.620e-8	2.5665	1.880
$\sigma_{1,m}^2$ [K^2]	9.780e-3	9.562e-3	8.680e-7	4.156e-7	2.0884	1.656
σ_2^2 [K^2]	1.900e-4	2.072e-4	3.700e-9	4.145e-9	0.8926	1.997
UA [W/K]	11.082	11.084	0.0115	0.0028	4.1290	0.107
CI [MJ/K]	2.2862	2.2837	1.469e-3	3.954e-4	3.7159	0.460

Table 7.2. *Results from estimation of $n_e = 50$ series, using 4-step predictions in the criterion. The mean autocorrelation of residuals is, $\bar{\rho}(1) = 0.712$.*

In Table 7.2 the results from estimation using a 4-step prediction-horizon in the criterion are shown. In this ideal case where the true model is contained by the model set there is no difference in the estimated parameters which is also seen from the tables, they are still unbiased. A problem, though, when "filtering" the residuals in this way is that the residuals are no longer white noise. From the tables we have that the mean autocorrelation of residuals are $\bar{\rho}(1) = 0.712$ when using the criterion based on 4-step predictions against $\bar{\rho}(1) = -0.001$ for 1-step predictions. When the residuals are autocorrelated a number of statistical tests for model validation are no longer valid. Another problem in Table 7.2 is that the uncertainties of the parameters are underestimated when the autocorrelation of residuals is not taken into account in the calculation of the uncertainties. Wahlberg & Ljung (1986) has shown, that asymptotically, it is only the prediction horizon itself, that affects the weighting in the frequency domain, and not how it is split up into sampling interval times number of predicted sampling instants. This means that, asymptotically, we could obtain the same results as using a k-step prediction, by using a proper anti-aliasing filter followed by a new sampling of the data and then estimate with a one-step prediction criterion. In this way one could avoid the autocorrelated residuals.

7.3 Selection of Input Sequence

In this section a Monte Carlo approach similar to that described in the previous section is used for choosing an optimal input signal among a number of candidates. The details of the study is presented in (Melgaard et al., 1992b). The model for this study is the same as in the previous section, but the sampling time has been fixed to one hour. For each candidate of input signal 100 sample paths of the system has been simulated. Then Fisher's information matrix is estimated on the basis of these sample paths. The determination of optimal inputs for the same system is discussed in (Sadegh, 1993), and the main results are briefly presented in the last part of this section.

Different test signals for the heating power have been investigated in order to compare their properties in relation to the different optimality criteria, described in Section 6.2. Most attention is paid to binary signals because of their simplicity (they are easy to implement). All binary signals switch between 0 W and 300 W and they all have equal power. The set of considered test signals is shown below. A step input and a number of PRBS sequences with different clock periods and orders have been considered. Also, a few signals containing two sinusoids, with the same total power as the binary signals, have been tested. PRBS sequences with increasing clock

Sequence	Description
prbs1	PRBS ($T_{prbs} = 1$ h, order $n = 9$)
prbs2	PRBS ($T_{prbs} = 2$ h, order $n = 8$)
prbs3	PRBS ($T_{prbs} = 5$ h, order $n = 7$)
prbs4	PRBS ($T_{prbs} = 8$ h, order $n = 6$)
step	step of period 252 h
sin1	sinusoids ($\tau_1 = 3$ h, $\tau_2 = 54$ h, power $1:1$)
sin2	sinusoids ($\tau_1 = 3$ h, $\tau_2 = 54$ h, power $1:2$)
sin3	sinusoids ($\tau_1 = 3$ h, $\tau_2 = 54$ h, power $1:9$)

Table 7.3. *The considered test signals.*

periods have been selected to examine the influence of increasing the period of the PRBS sequence compared to the sampling time. In other words, an optimal k in $T_{prbs} = k \cdot T_{sampl}$ is sought. The step signal is a very low frequency signal commonly used as the first test signal for identification. sin1 - sin3 consists of two sinusoids with periods $\tau_1 = 3$ hours and $\tau_2 = 54$ hours, which are close to the time constants of the system. The partition of the total power between the two sinusoids changes from $1:1$ to $1:9$, putting more weight on the low frequency sinusoid.

The results from the calculation of the optimality criteria, see Section 6.2.1, for the considered test signals are shown in Table 7.4. The D_s-*optimality* was calculated when all parameters except the noise terms are of interest, thus considering the noise parameters as nuisance. Within each column, representing an optimality criterion, the optimal test signal is the one that minimizes the value of the criterion. It is seen from Table 7.4 that,

Sequence	D	D_s	C	E
prbs1	-41.74 (4)	7.786 (7)	0.327 (4)	0.291 (4)
prbs2	-42.42 (3)	7.121 (4)	0.321 (3)	0.289 (3)
prbs3	-43.11 (1)	6.360 (1)	0.308 (1)	0.278 (1)
prbs4	-42.82 (2)	6.486 (2)	0.319 (2)	0.287 (2)
step	-40.94 (5)	8.760 (8)	0.451 (5)	0.426 (5)
sin1	-39.29 (7)	7.328 (5)	1.551 (8)	1.474 (8)
sin2	-39.71 (6)	7.087 (3)	1.304 (7)	1.232 (7)
sin3	-39.18 (8)	7.609 (6)	1.231 (6)	1.155 (6)

Table 7.4. *The value of the criteria of the standard measures of information from Section 6.2, for the different test sequences. The value of the criterion is given, and a ranging of the results in parentheses.*

according to all optimality criteria, the PRBS sequences have the best performance and prbs3 is the optimal choice. The results clearly point out that we gain from increasing the clock period of the PRBS sequence compared to the sampling period. In this case, an optimum is found for $T_{prbs} = 5T_{sampl}$.

When dealing with physical models, nonstandard measures may be of interest, since the parameters of greatest interest might not be directly entering into the system description, but some transformation of these parameters. When designing optimal input sequences, it is very important to take the transformation into account.

The UA-value and the CI-value are parameters that are probably of major interest in many applications of building performance. The UA-value is the overall thermal transmission coefficient between the inside air and the outdoor surface, and the CI-value is the internal heat capacity, defined as the amount of heat needed for raising the room air temperature by 1 K. These parameters are calculated as functions of the model parameters and we can use the precision of these characteristic numbers as the basis for an information measure instead of using the precision of the model parameters as before.

The characteristic parameters are given as a function, $\mathbf{f}(\theta)$, of the parameters of the model. Then Gauss' formula is used to approximate the information matrix in the domain of the characteristic parameters, cf. Section 6.2.2, i.e.,

$$\mathbf{M}_{F,f}^{-1}(\theta) = \left(\frac{\partial \mathbf{f}}{\partial \theta}\right)^T \mathbf{M}_F^{-1}(\theta) \left(\frac{\partial \mathbf{f}}{\partial \theta}\right) . \tag{7.47}$$

Then any of the standard measures of information can be applied to $\mathbf{M}_{F,f}(\theta)$. For the model specified previously, the characteristic parameters UA and CI are calculated as

$$\begin{pmatrix} \text{UA} \\ \text{CI} \end{pmatrix} = \begin{pmatrix} H_m H_i/(H_m + H_i) \\ C_i + H_i C_m/(H_m + H_i) \end{pmatrix} . \tag{7.48}$$

In Table 7.5 the results from using the these application-oriented measures are given. Both optimality with relation to UA- and CI-criterion is calculated as well as D-optimality for the vector of the physical characteristic numbers. The purpose of the physical measures is to focus on the UA- and CI-values. It is clearly seen from Table 7.5 that the step sequence now gives optimal information. Among the sinusoids, the sequence that has the most weight on the low-frequency part has the best performance, and among the PRBS signals the one with the largest clock period is performs best. In summary, the signals that have a major part of the variation at low frequencies are optimal for identification of the UA- and CI-values. This

Sequence	UA	CI	$D_{UA,CI}$
prbs1	0.3068 (8)	485.0 (8)	70.14 (8)
prbs2	0.2596 (7)	268.1 (7)	45.55 (7)
prbs3	0.1180 (3)	164.3 (6)	15.80 (5)
prbs4	0.0907 (2)	132.2 (4)	10.71 (3)
step	0.0374 (1)	67.04 (1)	2.507 (1)
sin1	0.1636 (6)	138.1 (5)	18.89 (6)
sin2	0.1596 (5)	94.78 (3)	13.52 (4)
sin3	0.1453 (4)	71.80 (2)	10.03 (2)

Table 7.5. *Results from calculation of the physical measures of optimality. The ranging of the results is given in parentheses.*

corresponds nicely to the fact that these values mostly affect the frequency response at lower frequencies.

The conclusion of the study is that signals that are commonly considered generally good input signals, which is the case for PRBS signals, may not be optimal when the purposivity of the model is taken into account by using some physical measures of optimality. At least the weighting of frequencies of the optimal input sequence can change remarkably by changing the criterion of optimality, in this case towards low frequencies.

In this section only the selection of input sequence among a number of alternatives has been considered. The Monte Carlo approach is easy to apply for complex models also when non-controllable inputs are present (the outdoor climate). Alternatively an analytical solution of the problem might be considered. This approach is discussed in Chapter 6. By estimating a frequency spectrum for the external surface temperature, it should be possible to find the analytical solution for the considered case. It will though, require some computations, compare with the example of optimal design for the scalar case of an embedded continuous time stochastic model with discrete time data in Section 6.5.

The design of D-optimal experiments for the system considered in this section is discussed in (Sadegh, 1993). In his study, though, the influence of the external climate, (T_e), is neglected. By assuming a large experiment time and a power constrained input power, the results of Section 6.3 apply. The design can be formulated as an optimization problem,

$$\begin{aligned}
\min_{\alpha_1,\ldots,\alpha_n,\omega_1,\ldots,\omega_n} \quad & f(\alpha_1,\ldots,\alpha_n,\omega_1,\ldots,\omega_n) \\
\sum_{i=1}^{n} \alpha_i &= 1 \\
0 \leq \alpha_i \leq 1 & \qquad i = 1,\ldots,n \\
0 \leq \omega_i \leq \infty & \qquad i = 1,\ldots,n
\end{aligned} \tag{7.49}$$

with the cost function

$$f(\alpha_1,\ldots,\alpha_n,\omega_1,\ldots,\omega_n) = -\log\det[\sum_{i=1}^{n} \alpha_i \bar{M}_F(\omega_i)] \tag{7.50}$$

where ω_i is the single frequency and α_i is the power proportion of that frequency. It turns out, from geometrical considerations, that an optimal design exists comprising not more than $n = 2$ sinusoids for this case. The optimal frequencies are $\omega_1 = 0h^{-1}$ and $\omega_2 = 0.5h^{-1}$, with power proportions $\alpha_1 = 0.813$ and $\alpha_2 = 0.187$ respectively. The optimal design is shown in Figure 7.10 below. The poles for the transfer function corresponding to the test cell model are

$$p_1 = -0.0184 \qquad p_2 = -0.3300\,, \tag{7.51}$$

which are also marked on Figure 7.10. It is seen that the rule of thumb, to excite a system at the frequencies close to the eigenvalues of the system is

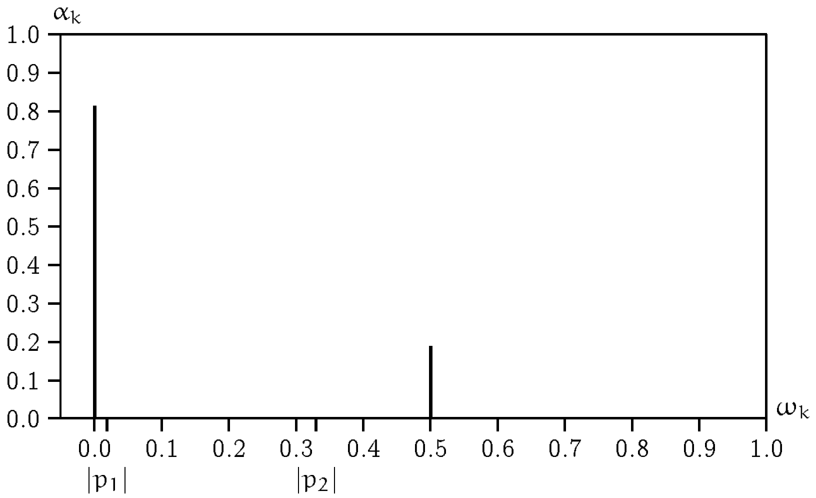


Figure 7.10. The D-optimal input power distribution for the test cell model. The optimal input signal comprises 2 sinusoids.

also valid here. There seem, though, to be some discrepancy between this optimal design and conclusions of the previous simulation study, see Table 7.4. There may be more reasons for this difference. The experiment length is about 9 times the largest time constant in the simulation study. For the optimal design, infinite experiment time is assumed. Another important reason for the difference may be the sensitivity of the 2 sinusoids input sequence to the specific choice of frequencies. It may seem that the sequence sin3 is quite close to the optimal design, whereas the performance of this sequence is not very good in the simulation study. The frequencies in sin3 are exactly at the poles of the system, while the frequencies of the optimal design are shifted towards lower and higher frequency, c.f. Figure 7.10.

7.4 Identification of Passive Solar Components tested in situ

The case that stimulated this work on physical modelling and grey-box estimation is related to a CEC research project on identification of thermal characteristics of building components tested in situ. In Figure 1 is shown a rough sketch of a test building used for testing the building components which is placed as the south wall of the building. In this case the wall to be tested consists of a lightweight insulated sandwich panel, provided with a double glazed window. Also in Figure 7.4 is mentioned the main

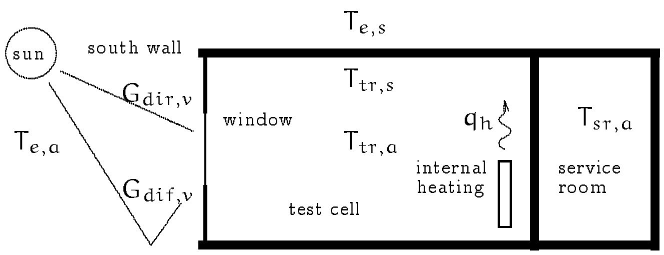


Figure 7.11. *A cross section of the test building with an installed south wall, here a lightweight component with a double glazed window in the middle.*

measured quantities, which includes $T_{tr,a}$, $T_{tr,s}$: internal air temperature and surface temperature, respectively, $T_{sr,a}$: the air temperature of the service room, $T_{e,s}$, $T_{e,a}$: the external surface and air temperature of the test building, respectively, q_h: the internal heat supply, and $G_{dif,v}$, $G_{dir,v}$: global vertical diffuse and direct solar radiation on south, respectively. The total vertical global radiation is $G_v = G_{dif,v} + G_{dir,v}$. The solar angle of incidence, θ_{inc}, is calculated from a physical model based on the ground location and the time of the year.

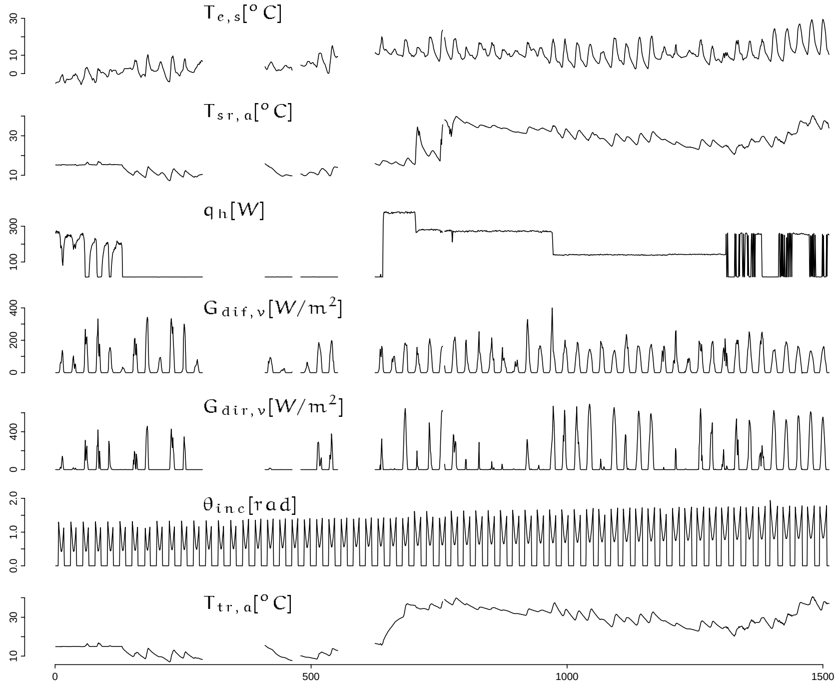


Figure 7.12. *Some of the data series from the experiment. The unit on the horizontal axis is hours. The holes in the series are missing observations due to different errors during data logging of the experiment.*

The procedure is first to perform a calibration experiment with a thick opaque insulation panel replacing the test component and subsequently identify the model of the test cell from this experiment. Then a second experiment is performed with the test component installed, and by including the information obtained from the calibration experiment it is possible to extract information about the test component separately. One of the questions to ask about this procedure is *how to incorporate the information from previous experiments in the estimation on the new data in a proper way.* One possibility, and the usual approach, is to fix the previously estimated parameters in the total model of the test cell and test component and estimate the remaining parameters. Due to the rather high correlation between some of the inputs to the model, mainly between the climate series, it is necessary that some of the parameters are fixed, in order for the system to be identifiable. The parameters to fix have been estimated from the calibration experiment, with an associated uncertainty. If the uncertainty is small there is no problem with this approach, but this is not always the case. Therefore a Bayesian approach is proposed, where the prior information is specified as a prior distribution function, and hence accounts for the associated uncertainty of the parameters. Since the parameters from the calibration experiment are ML estimates they are approximately normally distributed, see Section 3.3.2. Thus the prior distribution is determined uniquely from the estimated parameters and their associated covariance matrix. Then MAP estimates for the model are obtained by maximizing (3.52) as in Example 3.1.

The model of the test building is mainly build up as an R-C network, of heat resistors and heat capacitors by analogy with an electrical network. This is basically a linear model which is a lumped model of the equations for heat diffusion. The main structure of the model is shown in Figure 7.4. The top branch of the model in Figure 7.4 represents the east wall, west wall, floor and roof test cell. The middle branch is the partition wall between the service room and the test cell. In the calibration experiment the parameters of these to branches of the model are estimated. The south wall is then replaced by a calibration wall, which is highly insulated and homogeneous. Thus, the heat flow through this wall can be measured by a heat-flow meter during the calibration experiment. In the calibration

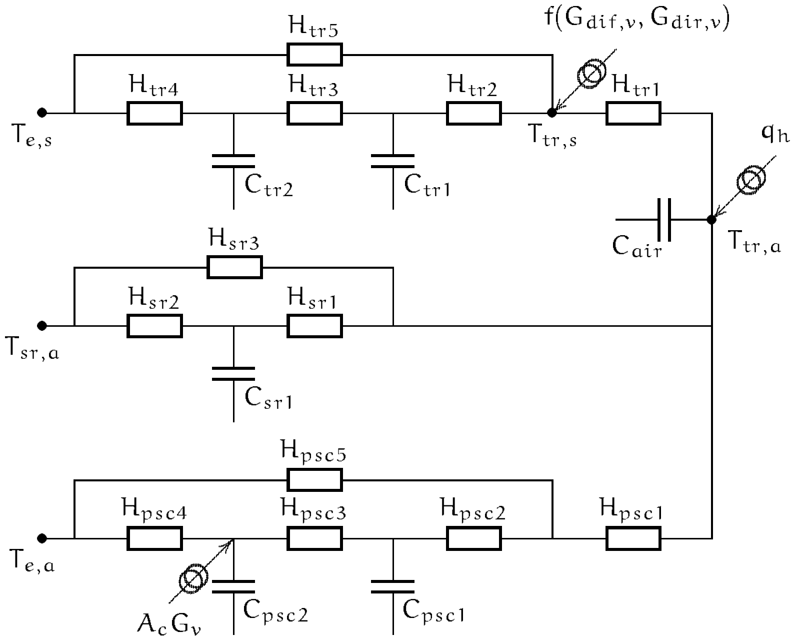


Figure 7.13. Network model of the test building with the installed test component.

experiment there is no solar radiation entering the internal test cell surface (the arrow on the top branch of the model). The parameters of these two branches of the model can then either be fixed, or included as a prior distribution function, when a MAP estimator is used for estimation based on the new experiment.

The bottom branch of the network model in Figure 7.4 describes the heat transfer through the test wall. The parameter H_{psc1} is fixed to some prior expected value, expressing the prior expected conductance between the middle of the test cell and the internal surface on the south wall. $f(G_{dif,v}, G_{dir,v})$ is a nonlinear function describing the solar transmittance through the window. The direct radiation is known to have a transmittance which is dependant upon the angle of incidence of the radiation, due to reflection by the window. A theoretical expression for the direct transmittance with relation to losses by reflections for different numbers of glass covers is used as the basis for estimating the nonlinear relation for a real window, which also includes a frame etc.

$$f(G_{dif,v}, G_{dir,v}) = A_{diff} G_{dif,v} + A_{dir} F(\theta_{inc}, \alpha_{eff}) G_{dir,v} \tag{7.52}$$

where A_{diff} and A_{dir} are constants (to be estimated). $F(\theta_{inc}, \alpha_{eff})$ is the theoretical function for the direct transmittance dependent on the angle of incidence and for α_{eff} number of glass covers. The expression is plotted in Figure 7.4. Theoretically the shape of the curve, in this case, should

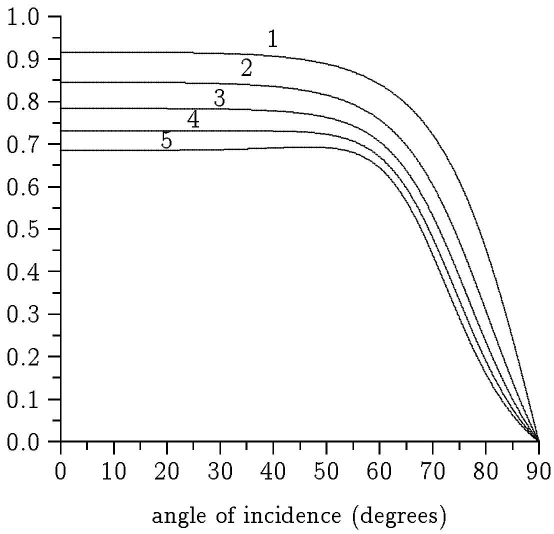


Figure 7.14. *Theoretical direct transmittance w.r.t. losses by reflections for different number of glass covers.*

correspond to double glazing, but due to unmodelled boundary conditions, (the frame etc.), the shape of the curve is estimated from the data. The estimated number, α_{eff}, may be considered as the effective number of glass covers, which can take any positive real number.

The nonlinear model is compared to a linear model, which was used earlier in the PASSYS project. For the linear model, the expression for the solar transmittance is given by

$$f(G_{dif,v}, G_{dir,v}) = A_{tot}(G_{dif,v} + G_{dir,v}) \ , \tag{7.53}$$

where A_{tot} is some constant.

In addition to the linear R-C network in Figure 7.4 and the nonlinear expression for the direct transmittance shown in Figure 7.4, the total model of the system also contains a model for the noise. The noise model accounts for the measurement noise on the inputs and outputs of the system, and un-modelled dynamics. The inputs to the model are

$$\mathbf{u}^T = (T_{e,s} \;\; T_{sr,a} \;\; T_{e,a} \;\; q_h \;\; G_{dif,v} \;\; G_{dir,v} \;\; \theta_{inc})^T \ . \tag{7.54}$$

The outputs of the model are

$$\mathbf{y}^T = (T_{tr,a} \;\; T_{tr,s})^T \ . \tag{7.55}$$

The whole model can be put in the form of (4.25) and (4.26), described in Section 4.1.4, with an additive noise term on the state vector and an additive measurement noise on the output equation. We are then able to estimate the parameters of the model, using e.g. CTLSM, see (Melgaard & Madsen, 1993).

7.4.1 Results

From the overall model a number of characteristic physical parameters are calculated, these include: $UA_{tr,e}$, steady state overall thermal transmission coefficient between test room and outdoor surfaces, $UA_{tr,sr}$, the steady state overall thermal transmission coefficient between test room and service room, UA_{psc}, steady state overall thermal transmission coefficient for the test component, gA_{psc}, steady state overall solar transmittance, or total solar heat gain factor of the test component, which is the ratio of heat entering the test-cell caused by solar radiation on the component, divided by the intensity of incident solar radiation on the component, $CI_{tr,e}$, internal test room heat capacity, i.e. the amount of heat which goes into the test room-envelope as a result of a change from one steady state situation to the same steady state situation except for the indoor temperature being raised by 1 K, $CI_{tr,sr}$, similar for the partition to service room, and CI_{psc}, similar for the test component.

Four situations of the estimation have been considered and compared. fixed/ML versus MAP/ML estimates have been compared, and models with or without the nonlinear function for the solar transmittance have been compared. The main results from the four situations are shown in Table 7.4.1 below. Beyond the estimated values of the characteristic parameters and their associated standard deviation in brackets, is also listed the mean and variance of the one-step prediction errors, for the internal surface temperature, $T_{tr,s}$, for the different cases.

It should be mentioned that in all cases the estimated physical characteristics are reasonable compared to the expected theoretical values, except for $UA_{tr,sr}$ in both fixed/ML cases, in which the values are too low compared to the theoretical values. The theoretical values, however, only covers one-

	linear		non-linear	
Parameter	fixed/ML	MAP/ML	fixed/ML	MAP/ML
$UA_{tr,e}$ [W/K]	7.826 (0.096)	8.265 (0.066)	7.826 (0.096)	8.294 (0.088)
$CI_{tr,e}$ [MJ/K]	2.045 (0.045)	1.976 (0.015)	2.043 (0.045)	2.000 (0.023)
$UA_{tr,sr}$ [W/K]	1.001 (0.148)	2.096 (0.237)	1.001 (0.148)	2.170 (0.242)
$CI_{tr,sr}$ [MJ/K]	0.274 (0.024)	0.350 (0.006)	0.274 (0.024)	0.355 (0.009)
UA_{psc} [W/K]	6.047 (0.044)	5.885 (0.055)	5.917 (0.049)	5.735 (0.094)
CI_{psc} [MJ/K]	0.136 (0.006)	0.153 (0.006)	0.116 (0.009)	0.119 (0.011)
gA_{psc} [m^2]	0.573 (0.006)	0.594 (0.008)	0.612 (0.004)	0.633 (0.005)
α_{eff}	- -	- -	3.391 (0.811)	3.401 (0.340)
mean p.e.	-0.0076	0.0016	-0.0054	0.0004
var p.e.	0.0151	0.0122	0.0131	0.0105

Table 7.6. *Main results from the estimations.*

dimensional heat loss; i.e. thermal bridges etc. are not taken into account. The gA values can not be compared directly, because the linear cases consider some mean value of the solar aperture, whereas in the non-linear cases it is an angle dependant function, (the numbers are specified for 25% diffuse radiation, and 45° angle of incidence for the direct radiation).

The difference in the estimated physical parameters from linear to non-linear model is small, except for the gA value; but the inclusion of the non-linear part has increased the capability of the model for prediction, the variance of the prediction error has decreased about 13 % for both the ML and MAP case. The larger difference is between the ML and MAP approach. In the pure ML approach, a number of the parameters have been fixed to the expected value estimated from the calibration experiment, whereas in the MAP approach, all parameters are estimated with a prior distribution determined by the calibration experiment. It is seen, that the physical characteristics which are mainly determined by the calibration experiment, i.e. $UA_{tr,e}$ and especially $UA_{tr,sr}$ have changed significantly from their prior expected value to their posterior expected value. This means, that despite the fact that the inputs of exterior climate variables to the model are rather high correlated, there is still enough information in the new experiment to change these physical characteristics. Especially for $UA_{tr,sr}$ it was only possible to obtain vague information from the calibration experiment, due to a badly excited input signal, $T_{sr,a}$, for that part of the model. This fact was also confirmed by visually plotting the inputs for the calibration experiment. In the new experiment this input signal has a better excitation. Still, though, this signal does not have a very good excitation, compare Figure 7.12, for large parts of the experiment $T_{sr,a} = T_{tr,a}$.

It has been shown that the proposed method and model are superior to the earlier approach for modelling this system. This includes the use of MAP estimation as a means of including the prior information about the system, and the use of a nonlinear model for the solar transmission through the window. A plot of the measured and simulated output from the model gives an indication of the performance of the model, see Figure 7.15. Even though the model seems to perform well with respect to simulated output, it is still not perfect, which is clearly seen from a cumulative periodogram of the residuals (the one-step prediction errors). A plot of the cumulative

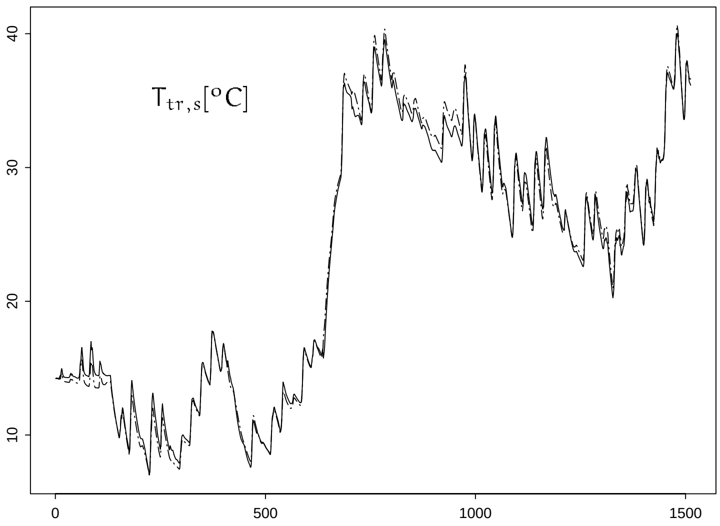


Figure 7.15. *Plot of* $T_{tr,s}$, *measured (solid) and simulated from the proposed model (dashed). The unit on the horizontal axis is hours.*

periodogram for the residuals of the two outputs of the model is given in Figure 7.16 below. From the cumulative periodogram it is seen that none of the residual sequences will pass a white noise test. This is especially true for $T_{tr,s}$. A closer look at the plot reveals that the curve for $T_{tr,s}$ has some well defined steps. These are located at the frequency of daily cycle and the higher harmonics of this frequency. This could indicate that the model of the transmitted solar radiation hitting the internal surface of the test cell (floor and wall) need further improvements. This may not be possible with the current measurement setup. Either more temperature sensors are needed at the inside surface of the test cell (where the sun hits the floor) or it should be realized that $T_{tr,s}$ can not be modelled by a one-dimensional model of the test cell, when the test wall has a window.

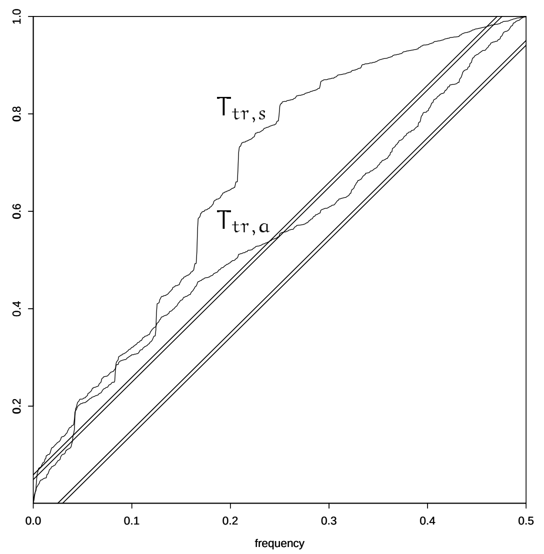


Figure 7.16. *The cumulative periodogram of the residuals of the two outputs of the model. The 95 % and 99 % confidence limits for a white noise test are depicted.*

7.4.2 conclusion

The incorporation of prior information for system identification has been considered twofold in this case. First by considering a model structure determined by a physical model of the system extended with a model of the noise. The strength of this approach is demonstrated by an example of estimation of physical parameters of building components, based on measurements from the system. The second way of including prior information is by the use of MAP estimation, where a prior distribution function of the parameters is specified and incorporated in the estimation. The example demonstrates that this approach is a better way of including prior information, than simply fixing certain parameters of a large model to their prior expected values.

The Bayesian approach makes a nice frame for combining information from previous experiments with new data. In this study it was demonstrated specifically for the parameters of the partition wall (the middle branch on Figure 7.4). In the calibration experiment the excitation of this part of the model was pour, but when the uncertain information is combined with the new data, the result is good also for this branch of the model. This also means that identifiability problems need not be very severe, because the (vague) information can still be combined with information from new experiments.

7.4.3 Future Experimental Setup

It has been realized that the current setup of the test cell with highly insulated walls results in a system with a large time constant from the heat transfer of the walls. Since the time constant of the internal heat dynamics is short, the test cell is a stiff system. This makes the identification of the system more complicated and a long experiment duration is necessary.

For future setup of experiments a modification of the test cell has been proposed. On all the inside surfaces an extra layer is mounted with heat flow meters and heating elements with a feedback control, thus working like a pseudo adiabatic shell. This setup has a number of advantages, by eliminating the long time constant of the test cell. The identification is more accurate and the experiment duration is much shortened (from 9 weeks to about 4 days).

7.5 Summary

In this chapter the use of physical modelling for identification within the PASSYS project has been demonstrated. The advantages compared to traditional testing of building components is that full scale passive solar components are tested under real, dynamic outdoor conditions and not under some artificial laboratory conditions.

Different aspects of the work of parameter estimation are discussed in the chapter. First the use of lumped parameter models as approximations of the heat diffusion equation is discussed. Then the estimation tool, CTLSM, is validated by estimating the parameters from simulated sequences. It is verified that the program gives reliable parameter values *and* reliable uncertainties of the parameters. Such a validation of the tool is very important before using it on real data. In this way numerical approximations etc. from the tool can be separated from modelling approximations applied to real data.

A Monte Carlo approach is used to estimate the information from a certain experiment. This approach is used to select the optimal input heating sequence for the test room. It is found that if all the parameters of the model is of equal interest then a certain PRBS sequence is the optimal input. If, on the other hand, only the global thermal characteristics, like UA and CI values are of interest then a very low frequency input, as a step function gives the maximum information. Results from optimal design of input sequence are also compared.

Finally CTLSM is used for the identification of a real building component, a light weight wall with a window. Different approaches are compared and a Bayesian approach is proposed as a means of combining the information from previous experiments with new data. Furthermore a nonlinear (angle dependent) model for the solar transmittance through the window is shown to perform better than the previous linear model.

Chapter 8

Case #2 Car Engine

Compact engine models often consist of a set of nonlinear differential equations which predict the time development of the mean value of the engine state variables and perhaps some other internal variable. Such models are sometimes called mean value engine models, see e.g. (Hendricks, 1992).

In this chapter the identification of a fuel flow submodel for a mean value engine model is described. The experiments are conducted on a 1.1L spark ignition (SI) Ford engine. The engine has a central fuel injection (CFI). The fuel flow submodel is of specific interest because it has a large influence on the fuel/air ratio during transient operation of the engine. Exact control of the fuel/air ratio during transient operation is important for controlling the pollutant emissions and fuel economy of the engine. A good model of the fuel flow is needed in order to implement a feed-forward control on the engine. This is much faster than feed-back control which is limited by an un-avoidable delay through the engine.

8.1 Introduction

In spite of the large interest in identifying physical parameters of engines there is very little work of this nature reported in the literature. This is mainly because an engine is a nonlinear system which is even difficult to describe physically and it is a very noisy control object which requires the use of statistical identification techniques. These difficulties are reflected in the rather sparse collection of literature references in this area and in the large variety of engine and engine model types treated in that which is available. The models identified range from very simple continuous or discrete transfer function models to linearized continuous mean value engine models. Mean value engine models are continuous dynamic models which predict the mean value of important engine variables several engine cycles. References to earlier work in this field are given in (Hendricks, 1992; Melgaard et al., 1990; Hendricks & Sorenson, 1990).

8.2 Model Formulation

In order to successfully estimate parameters in a model of a dynamic system, it is very important at the stage of experimental design to define the frequency ranges of the important dynamic engine subsystems. This is necessary, since for practical estimation it is not possible to estimate simultaneously time constants which differ too widely at the same time. In the table below, the characteristic frequency ranges of the most important SI engine subsystems are tabulated. For control and condition monitoring applications the fuel flow and rotational dynamic subsystems are the most important identification objects. This is because the fuel flow dy-

engine subsystem	subsystem bandwidth		
temperature changes	≈	0.01	Hz
rotational dynamics	≈	0.2	Hz
fuel flow dynamics	≈	2	Hz
manifold filling	≈	20-200	Hz
noise from crankshaft rotations	≈	20-80	Hz
noise pulses from piston	≈	30-170	Hz
noise from bearings, gear etc.	≈	400-600	Hz

Table 8.1. *The characteristic frequency ranges of the most important subsystems of a SI engine (Collacott, 1977, pp. 170–178).*

namics are of great significance for controlling the fuel/air ratio, ϕ, while the crank shaft dynamics determine an engine's drivability characteristics. The crank shaft dynamics also reflect the condition of the engine. In these experiments we consider only warmed up engines, i.e. temperature effects are ignored here.

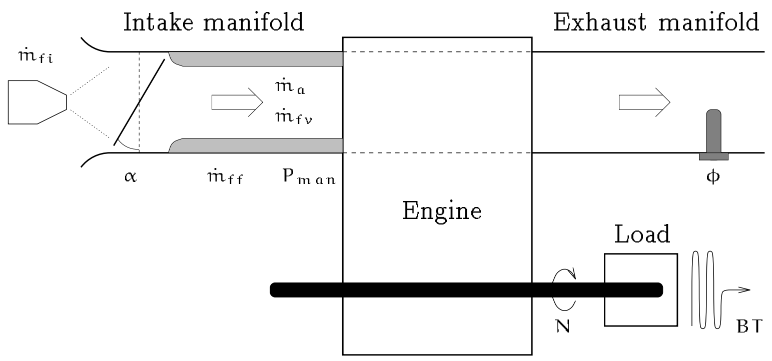


Figure 8.1. *Schematic block diagram of a CFI engine.*

In Figure 8.1 a schematic block diagram of the CFI engine is given. The relevant physical components and physical variables are indicated. These are the fuel injector, providing the injected fuel flow, $\dot{m}_{fi}$ [g/s]; the throttle plate, with an opening angle, α [deg]; the intake manifold, from where the manifold air pressure, P_{man} [kPa], is measured; the engine, where the energy contained in the air/fuel mixture by combustion is transformed into kinetic energy (and heat), the crank shaft speed is N [rpm]; the load, which is a dynamometer in this case, absorbs the kinetic energy and measures the brake torque, BT [Nm]; the exhaust manifold, at the end of which an oxygen sensor (= lambda sensor) is mounted to measure the fuel/air ratio, ϕ. The intake air mass flow, $\dot{m}_a$, is measured before the intake manifold. The fuel mass flows, $\dot{m}_{fv}$ and $\dot{m}_{ff}$, which denotes vaporized and fuel film respectively, are not measured directly, but they are part of the fuel flow model.

The fuel flow dynamic submodel for a CFI engine has been identified in the literature using classical identification techniques, e.g. (Aquino, 1981). For this study the model proposed by Aquino (1981) has been used as the basic identification object. The model is a semi-empirical representation of the behavior of the fuel film in the intake manifold. In this case the intake manifold is heated by the engine coolant to 80°C for a warmed up engine. A block diagram of the fuel flow submodel is shown in Figure 8.2. As indicated on the figure, the injected fuel mass flow, $\dot{m}_{fi}$, divides into two contributions: a vapor phase mass flow, $\dot{m}_{fv}$, and a liquid phase mass flow, $\dot{m}_{ff}$, which is the fuel film. The proportion of the fuel which goes into the fluid phase is $X \in [0, 1]$ while the remaining proportion $1 - X$ is entrapped in the air stream as vapor. The time constant which describes the dynamics of the entrapment process is expected to be of the same order as the manifold filling dynamics, and is considered to be very fast compared to the bandwidth of the fuel flow dynamics. The time constant, τ_{ff} describes the mean evaporation time for the fuel film from the intake

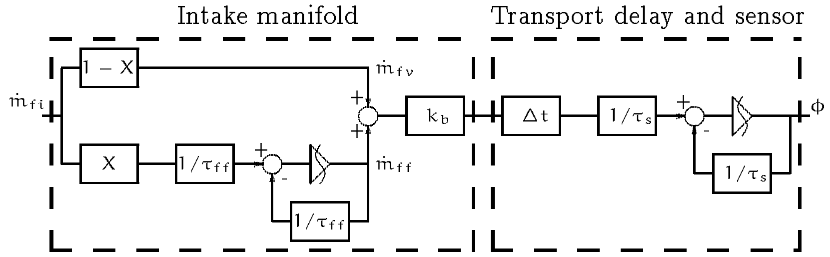


Figure 8.2. *The fuel flow model. The constant* k_b *is given by* $L_{th}/\dot{m}_a$*, where* $L_{th} = 14.67$ *is the mass ratio for stoichiometric mixture of the fuel and* $\dot{m}_a$ *is the air mass flow.*

manifold.

In order to complete the fueling dynamics submodel, a model for the dynamics of the lambda sensor must be given. A lambda sensor measures the air/fuel mass ratio normalized with the ratio at stoichiometric conditions: $\lambda = \dot{m}_a/(\dot{m}_f L_{th})$, where $L_{th} = 14.67$ is the mass ratio for stoichiometric mixture of the fuel. The fuel/air equivalence ratio, ϕ is given as the inverse of λ, i.e. $\phi = \lambda^{-1}$. The dynamics of the sensor is approximated by a first order low pass filter, with an expected time constant about 0.1 seconds according to the manual. There is a delay time included in the model, which is the time for the exhaust gases to pass through the engine and the exhaust valves and down the exhaust pipe to the lambda sensor. The term k_b, which is given by $k_b = L_{th}/\dot{m}_a$ accounts for the mixing of air and fuel in the intake manifold. Since the airflow is kept close to constant during the experiment, k_b is estimated as a constant value.

8.3 Measurement Setup and Experiment Design

The experiments were conducted on a four cylinder, four-stroke, 1.1L Ford CFI engine mounted on an eddy current dynamometer. The engine was fully equipped with sensors for all the relevant engine input and output variables, these are the ones mentioned in connection with Figure 8.1 in addition to $T_{m}an$, the temperature of manifold, which is heated by the coolant, and θ, the spark advance angle, measured in degrees BTDC (before top dead center). The air/fuel ratio was measured with a NGK Micro Oxivision MO-1000 Air/Fuel Ratio Meter with its own linear O_2 sensor. All measurements on the engine were logged by a PC based data acquisition system.

To keep the engine at the desired operating point during the experiment, the engine was given constant input biases: injected fuel flow, ignition spark angle BTDC and throttle angle. All experiments were conducted under open loop conditions. For the dynamic experiments, a PRBS perturbation was superimposed on the injected fuel flow signal. The minimum time between switching of the PRBS was chosen to $T_{prbs} = 0.5$s, and the number of stages was chosen as $n = 6$, then it should be possible to estimate time constants in the approximate interval 0.05 sec to 6 sec. In order to avoid aliasing effects, all the data logging channels are prefiltered with identical fourth order analogue filters, with a bandwidth of 20 Hz. The sampling frequency was selected as four times this bandwidth, i.e. $T_s = 0.0125$ sec. Each experiment has 3000 observations, which is the storage limit for the logging system. The intention is to select a number of operating points in a speed-load map of possible operating conditions for the engine. It should be remarked that an engine is difficult to operate manually in open-loop control, because only a limited deviation of the fuel/air ratio from stoichiometric mixture is possible, and the spark angle must be adjusted simultaneously when changing the operating conditions (speed and load) of the engine.

8.4 Results

The state space formulation of the fuel flow sub-model including process- and measurement noise is given by:

$$\begin{bmatrix} d\dot{m}_f(t) \\ d\dot{m}_{ff}(t) \end{bmatrix} = \begin{bmatrix} -1/\tau_s & 1/\tau_s \\ 0 & -1/\tau_{ff} \end{bmatrix} \begin{bmatrix} \dot{m}_f(t) \\ \dot{m}_{ff}(t) \end{bmatrix} dt + \begin{bmatrix} (1-X)/\tau_s \\ X/\tau_{ff} \end{bmatrix} \dot{m}_{fi}(t)\, dt + \begin{bmatrix} dw_f(t) \\ dw_{ff}(t) \end{bmatrix} \tag{8.1}$$

where $w(t)$ is assumed to be a Wiener process with incremental covariance matrix

$$\Sigma = \begin{bmatrix} \sigma_{1,f}^2 & 0 \\ 0 & \sigma_{1,ff}^2 \end{bmatrix} \tag{8.2}$$

and $\dot{m}_f(t)$ is the fuel flow as measured by the lambda-sensor (except for the pure time delay). The measurement equation is written

$$\phi(t+\Delta t) = [k_b \; 0] \begin{bmatrix} \dot{m}_f(t) \\ \dot{m}_{ff}(t) \end{bmatrix} + e(t) \tag{8.3}$$

where $e(t)$ is the measurement error, assumed to be normally distributed with zero mean and variance σ_2^2.

The experiment outlined in the previous section was carried out for different points in the normal engine operating region, for varying values of the throttle angle and load. The maximum likelihood method is used for estimating the parameters of the model. In Table 8.2 the results from estimation on the different experiments are summarized. The operating region is indicated by the the throttle angle in degrees and the engine speed in rpm. In all cases the engine is warmed up, and the intake manifold is heated by the coolant at 80°C. Before making any conclusions on the estimated parameters it should be verified that the proposed model actually fits the measured data. By analyzing the residuals from the model it is confirmed that the model fits the data nicely. The cumulative residual periodogram for one of the data sets is shown in Figure 8.3. It is seen that the residuals are close to white noise, and a Kolmogorov-Smirnov test for white noise can not be rejected on a 5 % level (the dashed lines). This is also the case for the other data sets in Table 8.2. A plot of the simulated fuel/air

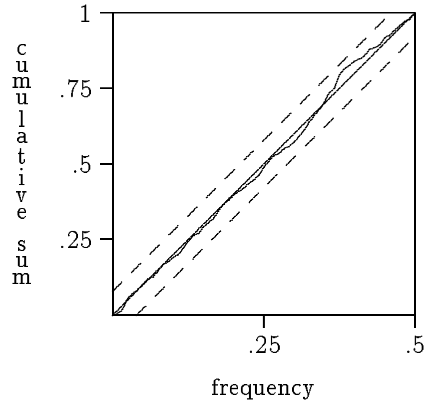


Figure 8.3. *Cumulative residual periodogram for the fuel flow model and data set s60g4, with 95 % confidence limits.*

equivalence ratio is compared with the measured ratio in Figure 8.4. This plot also indicates that the model is a good representation of the data. It is also seen that the measurements are quite noisy, and that the model is able to handle this fact.

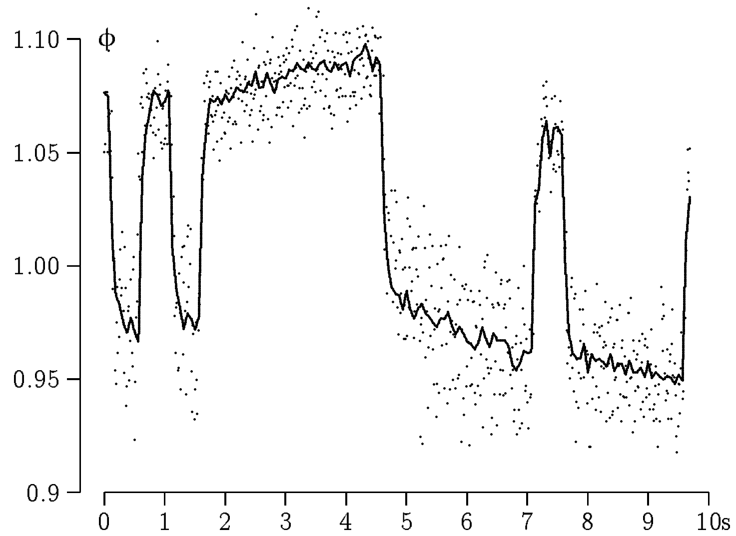


Figure 8.4. *The measured (dots) and simulated (line) fuel/air equivalence ratio, ϕ. Only the last part of the sequence is shown for the data set s60g4.*

Hence it has been verified that the proposed model is a good representation of the data for all the data sets. Then it will be investigated if and how the variation of the parameters of the model are correlated with the different engine characteristics at the different points of operation of the engine. In order to obtain some idea of the correlation of the variables a multivariate linear regression analysis was performed. The explanatory variables were, the throttle angle, α, the engine speed, N, the manifold pressure, P_{man}, and the air mass flow, $\dot{m}_a$. The response variables were, the sensor time constant, τ_s, the fuel film time constant, τ_{ff}, and the fraction of liquid fuel, X. A method known as *leaps and bounds* is used, which is a generalization of the *stepwise regression* methods, by operating with all possible subsets of the explanatory variables. The criterion is Mallow's C_p statistic, measuring the residual mean square against the mean square for the total regression, see e.g. (Weisberg, 1985).

First the variation of X, the fraction of liquid fuel, is investigated. From the regression analysis it is concluded that the best model is obtained by using N and P_{man} as explanatory variables. In order to visualize this correlation the variables are plotted in a 3-D plot by interpolating a surface on the observations, see Figure 8.5. The second best model is found for only one

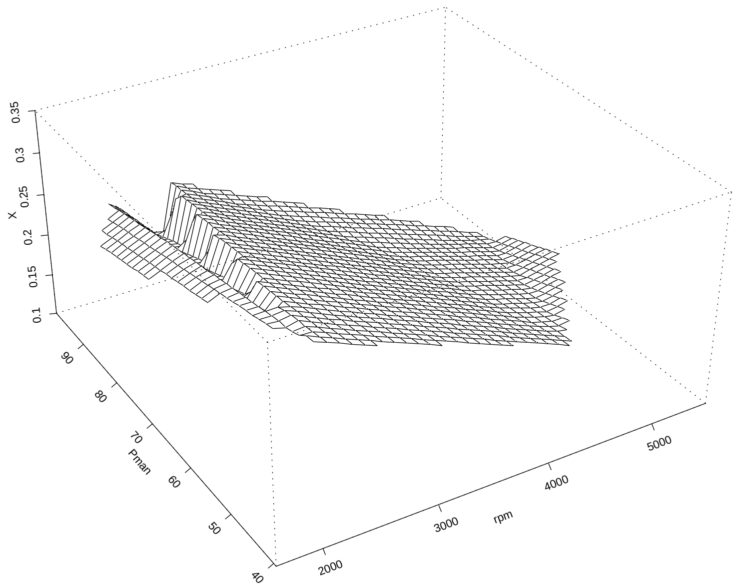


Figure 8.5. *Interpolated surface of* X *as function of* N *and* P_{man}.

explanatory variable, namely intake air mass flow. Figure 8.6 is a plot of X versus the airflow, the linear regression line is also drawn to show the correlation. In the same plot is also shown the data from (Melgaard et al., 1990), which were obtained from the same engine, but with a slightly different measurement setup. It is not expected that data from the two set of experiments are the same, but they are just plotted for comparison. It

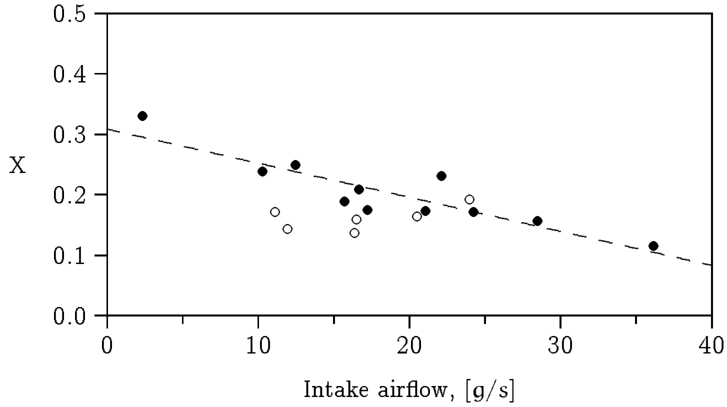


Figure 8.6. *The estimated fraction, X of the metered fuel flow entering the puddle in the intake manifold, plotted against the airflow. The manifold is heated by coolant at 80 °C. The empty circles are data from (Melgaard et al., 1990).*

is seen that X is decreasing as the air flow increases. Other authors, (Wu, Aquino, & Chou, 1983; Hendricks & Sorenson, 1990), have suggested that X is dependent upon the opening of the throttle angle, but with increasing X as the throttle angle increases. The theoretical argument is that there is a spray effect on the edge of the throttle plate when the opening is small. In Melgaard et al. (1990) this relation was also suggested, from estimation in a slightly different model, but based on measurements from the same engine as in this case. The data from the previous paper has been analyzed with the current model, and the resulting estimates are plotted as empty circles in the figures. When it is taken into account that there has been a recalibration of all instruments between the two sets of experiments and the different experimental conditions etc. it must be concluded that there is a good correspondence between the two batches. It is not expected, though, that data are the same, and the old data are not pooled in the analysis, but just plotted for a visual comparison. From the linear regression analysis it is concluded, that a model of X based on α alone is among the worst performing of possible models according to the criterion.

The observation of X being a decreasing function of the air flow is in good connection with empirical observations of other authors, e.g. (Heywood, 1989; Kay, 1978). Through transparent sections of the intake manifold Kay (1978) observed visually the presence of liquid fuel and puddles on the manifold walls under different operating conditions. They concluded that the largest amount of liquid fuel was present at high engine loads and low speeds, and at high engine speeds almost no fuel film could be detected. This corresponds nice to our observations, compare Figure 8.5.

The correlation between the estimated time constants and the operating region of the engine has also been investigated from the linear regression analysis. The sensor time constant, τ_s, is modelled best from α and N. A 3-D plot of the interpolated surface of τ_s as a function of α and N is shown in Figure 8.7 to indicate this correlation. It is seen that the sensor time constant is largest for low speed and small opening of the throttle plate. The second best model is with the sensor time constant as a function of the intake air mass flow alone. The correlation is seen in Figure 8.8. It is seen, that the time constant is decreasing when the air flow is increasing. From a physical point of view this is an expected behavior for the lambda sensor. The sensor is situated at the exhaust pipe, and the outer electrode of the sensor is exposed to the exhaust gases, which have to pass through a slotted shield to reach the electrode. It is obvious that this passing time will decrease when the flow rate is increased.

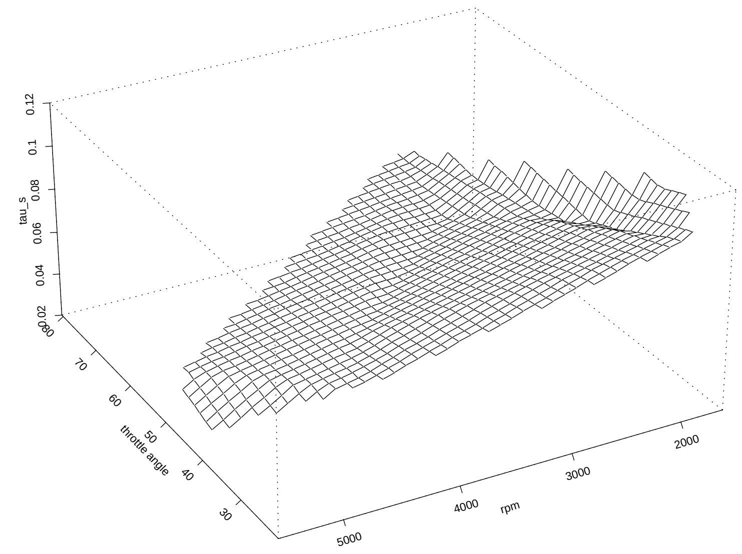


Figure 8.7. *Interpolated surface of* τ_s *as function of* α *and* N.

For the other time constant, τ_{ff} there seems not to be a very clear correlation with the operating conditions of the engine. There is, though, a tendency of decreasing value of this time constant for increasing engine speed, which is the best model from the linear regression analysis. A plot of τ_{ff} as a function of N together with the regression line is given in Figure 8.9. The mean of the estimated values of the fuel film time constant is

$$\bar{\tau}_{ff} = 1.22 \, \text{sec}$$

This can be compared to (Heywood, 1989, p. 319) where the time constant is said to be of the order of 1 second for typical manifold conditions.

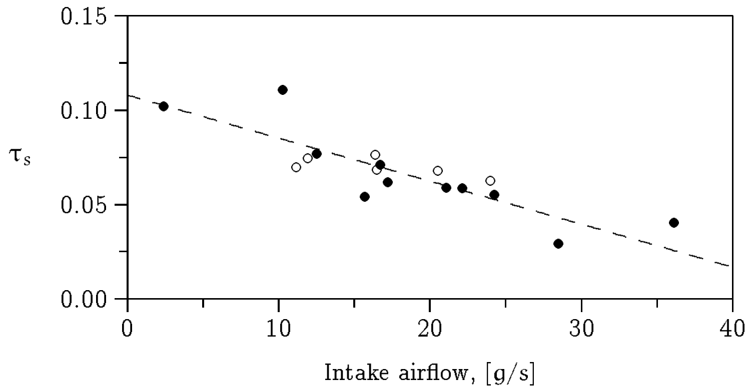


Figure 8.8. *The estimated time constant of the oxygen sensor τ_s in seconds plotted against the airflow.*

The estimated time delay, Δt, which represents the time lag from the intake ports and down the exhaust pipe to the lambda sensor is expected to be correlated with the mass flow through the engine. In Figure 8.10 the time lag is plotted versus the intake air mass flow. From physical considerations it is expected that the time delay can be modelled as a function proportional to the inverse of the air mass flow, which is indicated by the dashed line in the figure.

All the results from the linear multivariate regression analysis shown in the plots in this study have significant parameters in their models. Only the plots are given here, and not the individual values of the parameters. The number of experiments in the possible operating region of the engine is still to small to give the exact functional relations between the variables, and therefore only the linear correlation between the variables is considered in this study.

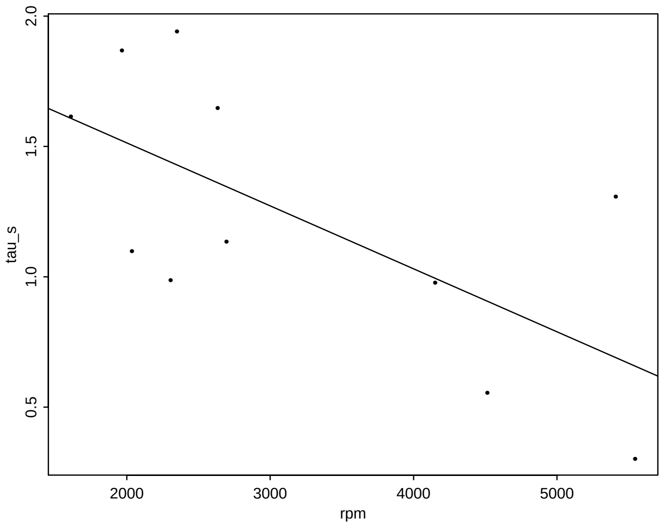


Figure 8.9. *Linear regression of* τ_{ff} *as a function of* N.

It can be seen from Table 8.2 that the variances of the noise terms in the model have some variation between the different experiments. When looking at the plots of the signals from the different experiments this fact becomes quite clear. The shape of the signals is different from one experiment to the other due to the different operating conditions of the engine. High speed will give a different noise in the signals than low speed etc. Furthermore the engine is very sensitive to the correct settling of the ignition spark angle. If this is not correct, the engine may be "knocking", which cause a large noise on the signals.

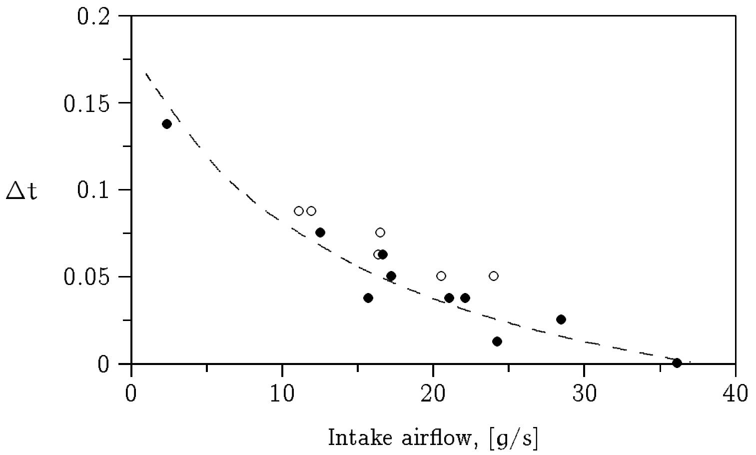


Figure 8.10. *The estimated flow transport delay, Δt from the injector to the oxygen sensor in seconds plotted against the air mass flow.*

8.5 Conclusions

In this study, the methods for grey-box identification have been used to identify an important submodel of an SI engine. The fuel flow model is of great importance especially for feed-forward control purposes because it has a large influence on the air/fuel ratio at transient operation of the engine. It has been possible to estimate the parameters of the model from measurements on the 1.1L CFI engine in a wide range of the realistic operation conditions for a warmed up engine.

Using these statistical methods it has been possible to verify some expected behavior of the model and getting insight to some more or less unknown behavior, e.g. the variation of liquid fuel fraction, X on a speed-load map of the engine. It was also possible by this method, to take into account the sometimes large and different levels of noise at the different operating conditions of the engine.

file	throttle	rpm	$\Delta t[s]$	$\tau_s[s]$	$\tau_{ff}[s]$	X	$k_b[s/g]$	$\sigma^2_{1,f}$	$\sigma^2_{1,ff}$	σ^2_2
s20g2	20.5	1970	0.138	0.102	1.87	0.330	2.41	0.80e-3	0.25e-5	0.44e-10
				(0.0048)	(0.296)	(0.014)	(0.071)	(0.90e-4)	(0.29e-5)	(0.12e-9)
s30g1	30.1	4520	0.038	0.054	0.56	0.188	0.80	0.22e-2	0.14e-7	0.71e-4
				(0.0020)	(0.072)	(0.011)	(0.008)	(0.96e-3)	(0.60e-8)	(0.16e-4)
s30g4	30.8	1610	0.038	0.111	1.62	0.238	0.89	0.22e-7	0.36e-5	0.73e-4
				(0.0058)	(0.436)	(0.027)	(0.031)	(0.66e-7)	(0.36e-5)	(0.72e-5)
s40g2	40.8	5550	0.025	0.029	0.30	0.156	0.41	0.28e-6	0.40e-7	0.31e-2
				(0.0069)	(0.173)	(0.036)	(0.016)	(0.29e-5)	(0.11e-6)	(0.19e-3)
s40g3	41.1	4150	0.013	0.055	0.98	0.171	0.45	0.39e-1	0.45e-4	0.30e-8
				(0.0015)	(0.138)	(0.014)	(0.007)	(0.21e-2)	(0.40e-4)	(0.43e-9)
s40g4	42.7	2040	0.075	0.077	1.10	0.249	0.59	0.17e-2	0.13e-3	0.60e-4
				(0.0040)	(0.266)	(0.033)	(0.026)	(0.94e-3)	(0.58e-4)	(0.11e-4)
s50g3	51.1	5410	0.000	0.040	1.31	0.114	0.38	0.67e-7	0.90e-4	0.19e-3
				(0.0021)	(0.435)	(0.025)	(0.010)	(0.54e-5)	(0.49e-4)	(0.12e-4)
s50g4	51.1	2310	0.063	0.071	0.99	0.207	0.47	0.58e-2	0.39e-4	0.36e-4
				(0.0038)	(0.240)	(0.023)	(0.012)	(0.16e-2)	(0.40e-4)	(0.99e-5)
s60g4	61.0	2630	0.038	0.059	1.65	0.172	0.40	0.47e-2	0.13e-7	0.84e-4
				(0.0027)	(0.374)	(0.020)	(0.009)	(0.27e-2)	(0.78e-7)	(0.13e-4)
s70g4	71.0	2700	0.038	0.059	1.14	0.231	0.40	0.27e-3	0.24e-3	0.68e-4
				(0.0028)	(0.274)	(0.031)	(0.016)	(0.73e-3)	(0.87e-4)	(0.57e-5)
s80g4	80.4	2350	0.050	0.061	1.94	0.173	0.43	0.86e-2	0.25e-3	0.61e-4
				(0.0038)	(0.836)	(0.067)	(0.033)	(0.34e-2)	(0.12e-3)	(0.15e-4)

Table 8.2. *Summarize of the experiments. The numbers in brackets are the corresponding standard deviations. The units of the two process variances are* $[g^2/s^2]$.

file	P_{man}[kPa]	$\dot{m}_a$[g/s]	$\dot{m}_{fi}$[g/s]	ϕ	BT[Nm]
s20g2	39.49	6.89	0.4385	0.9314	17.43
s30g1	45.67	17.07	1.1072	0.9498	14.04
s30g4	83.56	12.56	0.8546	0.9962	59.73
s40g2	70.26	33.09	2.3097	1.0220	28.54
s40g3	77.15	29.00	1.9932	1.0062	48.55
s40g4	93.25	18.03	1.2428	1.0092	71.85
s50g3	85.36	36.48	2.4197	0.9710	47.11
s50g4	95.58	19.69	1.3654	1.0154	73.01
s60g4	97.08	22.91	1.5819	1.0108	73.88
s70g4	98.06	23.48	1.6209	1.0107	74.72
s80g4	98.45	21.92	1.5064	1.0060	72.09

Table 8.3. *Summarize of the experiments.*

Chapter 9

Conclusion

In this thesis a number of aspects of identification of physical models have been considered.

It is argued for that physical models often can be formulated inside the framework of stochastic differential equations. By this formulation of the dynamical system, the physical understanding of the model and its parameters is preserved. At the same time the inevitable noise of a real world system may be included in the model description.

This approach for modeling of dynamical systems has some advantages compared to traditional black box modeling. The a priori knowledge is more easy incorporated in the model identification in terms of the model structure and a priori knowledge of the parameters. The physical knowledge is also used in the validation of the estimated model.

The thesis describes the different steps to be taken for identification of dynamical systems, when prior information is available and included in the identification:

- Design of the experiment. There are different levels for including the prior knowledge in the design.
- Choose the estimator for the parameters of the model. Mainly the maximum likelihood method has been considered, but the MAP estimator can also be useful for combining prior information with data.
- Define the class of model structures to consider, and select the corresponding state filter, which will produce the residuals for the estimator. (Unfortunately only a limited number of state filters are implemented at the time of writing this).
- Model validation. Many of the methods are closely related to the validation techniques for black box models, but the prior information about model structure and parameters is also included in the validation of the model.

The main drawback of this approach is the computational efforts involved. The estimation of parameters of stochastic differential equations from discrete observations is not a simple task, not even in the linear case. A thorough implementation of the numerical calculations involved is necessary, in particular the implementation of the predictor is important.

The methods of physical modeling have been applied to two different case studies. The first case is the determination of global thermal characteristics for building components, tested in situ.

First the use of lumped parameter models for approximating the heat diffusion equation is discussed. Then a way of validating the estimation tool is proposed, by which it is possible to test the bias and uncertainty of the estimated parameters. Such a validation of the tool is important, especially when it is going to be applied for estimation of lumped parameter models of buildings, because they are often described as a stiff system. When estimating stiff systems it is important that the numerics of the program is handled properly.

Then a Monte Carlo method is used for estimating the information matrix for a given choice of input sequence. The approach is used to select the optimal input among a number of candidates, from a given criterion based on the information matrix. It is concluded that the weighting of the importance of the individual model parameters has a large influence on the optimal choice of input sequence. If all the parameters of the model are equally weighted, then a certain PRBS is the optimal choice. If only the global thermal characteristics is important then a low frequency input, like a step input, is the optimal choice.

Finally the global thermal parameters are estimated for a light weight wall with a double glazed window. The use of a Bayesian approach for combining the previous information with new data is proposed as an alternative to the earlier approach of fixing parameters to their expected mean. The estimated parameters are found to be in good correspondence with expected theoretical values.

The second case study is the modeling of an important process of a spark ignition engine. A phenomenological model of the fuel flow including a fuel film in the intake manifold has been proposed in the literature. A good model of the fuel flow is important for controlling the air/fuel ratio under transient conditions. In particular when a feed-forward control should be applied. Although it appears that fuel films may have an appreciable influence on engine operation, attempts to quantitatively characterize that influence apparently have not been undertaken. By using stochastic differential equations as the basis for the modeling, it has been possible to use the proposed phenomenological model combined with a model of the noise, which is indeed present for operating engines. It is concluded that the proposed phenomenological model (combined with a model of the sensor dynamics) fits nicely to the measured data. The estimated models have gained new information about the fuel flow behavior at different operating conditions of the engine.

Appendix A

Implementation and Numerics

All practitioners have realized that the theoretical development of an algorithm is quite a different thing than the implementation of the algorithm on a computer. Important problems arise due to the finite arithmetic in the computer, which introduces a rounding error for every calculation with real numbers. In certain cases these rounding errors will accumulate so the solution could be misleading.

Even when a certain algorithm is chosen, there may be many algebraically equivalent ways of performing the calculations. The different ways of organizing the computations may have a substantial influence on the numerical properties of the algorithm, such as:

- computing time,
- memory requirement,
- numerical accuracy and stability related to rounding errors,

- programming effort.

A software tool called CTLSM has been developed for estimation of parameters in linear or non-linear stochastic differential equations, based on discrete time measurements, see also (Melgaard & Madsen, 1993). A large part of the important numerical considerations involved in this implementation is discussed in this appendix in the following sections.

A.1 Matrix Computations

A general class of problems, which is closely related to both the Kalman filtering algorithm and the optimization, is to find the unique matrix $\mathbf{X}$ which solves the least squares problem of minimizing

$$||\mathbf{B} - \mathbf{A}\mathbf{X}||_2 \tag{A.1}$$

where $\mathbf{A}$ is a real $m \times n$ matrix of rank n, $m \geq n$, and $\mathbf{B}$ is a $m \times l$ real vector.

Focus on an algorithm to solve (A.1) that minimizes the error on the solution $\mathbf{X}$ due to floating point operations, and of course with an eye on the computer time and memory, as mentioned above.

The condition number of a matrix is a measure of the sensitivity of matrix calculations to element perturbation. The condition number for the $\mathbf{A}$-matrix, corresponding to the L_2-norm, is defined as

$$\kappa(\mathbf{A}) = \sqrt{\max \lambda(\mathbf{A}^T\mathbf{A}) / \min \lambda(\mathbf{A}^T\mathbf{A})} \tag{A.2}$$

where $\lambda(\cdot)$ denotes eigenvalue. The solution to (A.1) is $\mathbf{X} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{B}$ when $\mathbf{A}$ has full rank. So generally when using t-digit binary arithmetic, it is not possible to obtain even an approximate solution to (A.1) unless

$$\kappa(\mathbf{A}) \leq 2^{t/2} \tag{A.3}$$

since $\kappa(\mathbf{A}^T\mathbf{A}) = \kappa^2(\mathbf{A})$ see e.g. (Björck, 1967). When $\mathbf{A}$ is $n \times n$ and has full rank the solution could be written $\mathbf{X} = \mathbf{A}^{-1}\mathbf{B}$ and the limit corresponding to (A.3) would now be

$$\kappa(\mathbf{A}) \leq 2^{t} \tag{A.4}$$

If further $\mathbf{A}$ is positive definite (PD), it is possible to perform a factorization $\mathbf{A} = \mathbf{S}\mathbf{S}^T$, where $\mathbf{S}$ is a square root of $\mathbf{A}$, see e.g. (Bierman, 1977). In the latter case the solution can be written $\mathbf{X} = \mathbf{S}^{-T}\mathbf{S}^{-1}\mathbf{B}$, and in this case the limit on $\mathbf{A}$ would be

$$\kappa(\mathbf{A}) \leq 2^{2t} \tag{A.5}$$

The accuracy of the solution, of course depends on the actual method used. One conclusion to these considerations is, that it is important to be aware of the type of the $\mathbf{A}$-matrix and to use an algorithm fit for that type. If we e.g. are dealing with a PD matrix and want to find its inverse through factorization, we will achieve accuracies that are comparable with those given by a general method (i.e. no factorization) that has twice the numerical precision of the factorization algorithm.

Methods for solving (A.1) depending on the specific type of $\mathbf{A}$, have been widely discussed in the literature, e.g.

1. $\mathbf{A}$ is general $m \times n$: *Modified Gram-Schmidt* orthogonalization is a good method, see (Björck, 1967; Jordan, 1968).

2. $\mathbf{A}$ is square $n \times n$: then *Gaussian elimination* is used (Jordan, 1968).

3. $\mathbf{A}$ is $n \times n$ and PD : methods based on *Cholesky decomposition* have good performance, e.g. (Bierman, 1977).

PD matrices are important items in both the optimization algorithm and the Kalman filter, as Hessian and covariance respectively, so PD matrices will be the subject for the next section.

A.1.1 PD matrices and Cholesky decomposition

If $\mathbf{A}$ is PD it can be decomposed as $\mathbf{A} = \mathbf{S}\mathbf{S}^\top$ where $\mathbf{S}$ is a nonsingular matrix. It is possible to constrain the matrix $\mathbf{S}$ to be triangular. Then the decomposition is the Cholesky decomposition. A popular factorization is known as $\mathbf{L}\mathbf{D}\mathbf{L}^\top$ *factorization*, also called square root free Cholesky decomposition, see e.g. (Fletcher & Powell, 1974). In this case $\mathbf{A}$ is written as

$$\mathbf{A} = \mathbf{L}\mathbf{D}_a\mathbf{L}^\top \tag{A.6}$$

where $\mathbf{L}$ is unit lower triangular, and $\mathbf{D}_a$ is diagonal. This factorization avoids scalar square roots unlike the traditional Cholesky decomposition.

A.1.2 Updating of LDL^T Factorizations

A problem with applications to both the Kalman filter update and recursions in quasi-Newton methods is how to modify a $\mathbf{LDL}^\mathsf{T}$ factorization.

Assume that an $n \times n$ PD matrix $\mathbf{A}$ is given in factorized form as (A.6), we want to compute the modified matrix

$$\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{G}\mathbf{D}_g\mathbf{G}^\mathsf{T} \tag{A.7}$$

where $\tilde{\mathbf{A}}$ is known from other considerations to be PD, and $\mathbf{D}_g$ is a diagonal matrix. Thus it is necessary to compute a unit lower triangular matrix $\tilde{\mathbf{L}}$ and a diagonal matrix $\tilde{\mathbf{D}}$ with $\tilde{d}_i > 0$ such that

$$\tilde{\mathbf{A}} = \tilde{\mathbf{L}}\tilde{\mathbf{D}}\tilde{\mathbf{L}}^\mathsf{T} \tag{A.8}$$

Different authors have treated this problem, and some relevant references are found in e.g. (Fletcher & Powell, 1974) and (Thornton & Bierman, 1980). It can be shown that the solution to (A.7) is equivalent to performing some orthogonalization methods normally used for the least-squares problem (A.1). The basic idea for solving (A.1) is to compute the factorization $\mathbf{A} = \mathbf{Q}\mathbf{R}$, where $\mathbf{Q}$ is orthogonal i.e. $\mathbf{Q}^\mathsf{T}\mathbf{Q} = \mathrm{I}$ and $\mathbf{R}$ is upper triangular. Then minimizing (A.1) is equivalent to minimizing $||\mathbf{Q}^\mathsf{T}\mathbf{B} - \mathbf{R}\mathbf{X}||_2$, which is easily solved by back-substitution. Some wellknown algorithms of this type are based on Householder, modified Givens or modified Gram-Schmidt orthogonalization, see (Golub & van Loan, 1983). Algorithms based on these transformations have comparable accuracy and numerical stability, so differences are basically in the number of computations and memory required, see (Björck, 1967; Jordan, 1968; Lawson & Hanson, 1974).

A method for solving rank one updates, that is (A.7) with $\mathbf{D}_g$ being a scalar, using modified Givens transformation is very thoroughly described in (Fletcher & Powell, 1974) and the method is implemented as a FORTRAN subroutine, code MC11A, which is in the Harwell Subroutine Library.

A method for solving (A.7) for a general $\mathbf{D}_g$ matrix, using modified Gram-Schmidt is described in (Thornton & Bierman, 1980). Their idea is to set

$$\boldsymbol{W} = [\mathbf{L}|\mathbf{G}] = \begin{bmatrix} \boldsymbol{w}_1 \\ \vdots \\ \boldsymbol{w}_n \end{bmatrix} \quad , \qquad \mathbf{D} = \mathrm{diag}(\mathbf{D}_a, \mathbf{D}_g) \tag{A.9}$$

and to view the rows of $\boldsymbol{W}$ as elements of a weighted inner product vector space, with an inner product defined as

$$\langle \boldsymbol{w}_i, \boldsymbol{w}_j \rangle = \boldsymbol{w}_i \mathbf{D} \boldsymbol{w}_j^T \tag{A.10}$$

One can then apply Gram-Schmidt orthogonalization to the row vectors of $\boldsymbol{W}$ and obtain an array $\tilde{\boldsymbol{W}}$, such that $\boldsymbol{W} = \tilde{\mathbf{L}}\tilde{\boldsymbol{W}}$, $\tilde{\mathbf{L}}$ is unit lower triangular and

$$\langle \tilde{\boldsymbol{w}}_i, \tilde{\boldsymbol{w}}_j \rangle = \tilde{\mathbf{D}}_i \delta_{ij} \tag{A.11}$$

where $\tilde{\boldsymbol{w}}_i$ and $\tilde{\boldsymbol{w}}_j$ are row vectors of $\tilde{\boldsymbol{W}}$, and δ_{ij} is the Kronecker delta. The result of the construction is that

$$\begin{aligned} \tilde{\mathbf{A}} &= \boldsymbol{W}\mathbf{D}\boldsymbol{W}^T \\ &= (\tilde{\mathbf{L}}\tilde{\boldsymbol{W}})\mathbf{D}(\tilde{\mathbf{L}}\tilde{\boldsymbol{W}})^T \\ &= \tilde{\mathbf{L}}(\tilde{\boldsymbol{W}}\mathbf{D}\tilde{\boldsymbol{W}}^T)\tilde{\mathbf{L}}^T \\ &= \tilde{\mathbf{L}}\tilde{\mathbf{D}}\tilde{\mathbf{L}}^T \end{aligned} \tag{A.12}$$

where $\tilde{\mathbf{D}} = \text{diag}(\tilde{\mathbf{D}}_1, \cdots, \tilde{\mathbf{D}}_n)$. Thus $\tilde{\mathbf{L}}$ is the transformation of $\boldsymbol{W}$ to "$\mathbf{D}$-orthogonal" coordinates. The construction of $\tilde{\mathbf{L}}$ and $\tilde{\mathbf{D}}$ via a modified Gram-Schmidt orthogonalization are summarized below

$\mathbf{LDL}'$-*updating algorithm*

$$
\left.\begin{array}{l}
\tilde{\mathbf{D}}_j = \langle \boldsymbol{w}_j^{(j)}, \boldsymbol{w}_j^{(j)} \rangle \\
\left.\begin{array}{l}
\tilde{\mathbf{L}}_{i,j} = \langle \boldsymbol{w}_i^{(j)}, \boldsymbol{w}_j^{(j)} \rangle / \tilde{\mathbf{D}}_j \\
\boldsymbol{w}_i^{(j+1)} = \boldsymbol{w}_i^{(j)} - \tilde{\mathbf{L}}_{i,j} \boldsymbol{w}_j^{(j)}
\end{array}\right\} \text{ i = j+1, } \cdots \text{, n}
\end{array}\right\} \text{ j = 1, } \cdots \text{, n-1}
$$

$$
\tilde{\mathbf{D}}_n = \langle \boldsymbol{w}_n^{(n)}, \boldsymbol{w}_n^{(n)} \rangle
$$

where the $\boldsymbol{w}$-superscripts indicate that the original $\boldsymbol{w}$ vectors are written over by the new ones in each iteration.

The modified Gram-Schmidt orthogonalization is described by (Lawson & Hanson, 1974); and (Björck, 1967) and (Jordan, 1968) have investigated its numerical characteristics. Their studies establish that the computed $\tilde{\boldsymbol{w}}$ vectors are nearly orthogonal, and that the $\tilde{\mathbf{L}}$ basis transformation coefficients and the $\tilde{\mathbf{D}}$ vector are relatively insensitive to the effects of computer round-off. Their results establish that the modified Gram-Schmidt algorithm has accuracy comparable to that attained using Givens or Householder transformation methods.

A.2 Optimization

A crucial point in any statistical analysis based on the maximum likelihood method is concerned with the actual maximization of the likelihood function. In our case an explicit method does not exist, and the maximization has to be found using a numerical method of iterative character.

The problem can also be formulated as to find the minimizing point of a nonlinear function $f : \mathbb{R}^n \rightarrow \mathbb{R}$

$$\min_x f(\mathbf{x}) \tag{A.13}$$

Among optimization methods, the Newton-Raphson's method is very effective when attainable. A minimum of f is found where $\mathbf{g} = \partial f/\partial \mathbf{x} = 0$. The method is based on a Taylor expansion of $\mathbf{g}$ to first order

$$\mathbf{g}(\mathbf{x}_n + \mathbf{h}) = \mathbf{g}(\mathbf{x}_n) + \frac{\partial \mathbf{g}}{\partial \mathbf{x}}(\mathbf{x}_n)\mathbf{h} + o(\mathbf{h}) \tag{A.14}$$

Putting $\mathbf{g}(\mathbf{x}_n + \mathbf{h}) = 0$ and neglecting $o(\mathbf{h})$ the algorithm takes the form

$$\mathbf{s}_n = -\mathbf{H}(\mathbf{x}_n)^{-1}\mathbf{g}(\mathbf{x}_n) \tag{A.15}$$

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \mathbf{s}_n \tag{A.16}$$

where the Hessian $\mathbf{H} = \partial^2 f/\partial \mathbf{x}^2$ in the regular case is positive definite. Since, in our case, we are not able to provide the optimization procedure with the first and second partial derivatives of $f(\mathbf{x})$, these will have to be approximated by the algorithm. We have to use finite-difference approximation to the first derivative g in (A.16), and we will use a secant approximation $\mathbf{B}_n$ to the Hessian. The secant approximation is more effective and robust than a finite-difference Hessian in the optimization. This class of secant methods are called *quasi-Newton*, and the most successful seems to be the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method for iterative Hessian approximation combined with soft line search (Dennis & Schnabel, 1983). The implementation of the minimization algorithm is an extended and modified version of the subroutine code VA13CD from the Harwell Library (1989).

A.2.1 Finite-difference derivatives

In order to estimate the gradient, an obvious method is the *forward difference* approximation

$$g_i(x_n) \simeq \frac{f(x_n + h_i e_i) - f(x_n)}{h_i} \quad , \quad i = 1, \cdots, n \tag{A.17}$$

where e_i is the ith basis vector, and the approximation having the error $o(h_i)$ in the ith component. Although this forward difference approximation in general is accurate enough sometimes *central differences* are useful because of better estimates

$$g_i(x_n) \simeq \frac{f(x_n + h_i e_i) - f(x_n - h_i e_i)}{2h_i} \quad , \quad i = 1, \cdots, n \tag{A.18}$$

which has error $o(h_i^2)$. The gradient computed with central differences is more accurate than the gradient based on forward differences, but it requires $2n$ rather than n evaluations of f assuming that $f(x_n)$ is already available.

The optimal choice of step size for the well-scaled case is for the forward differences

$$h_i = \eta^{1/2} x_i \tag{A.19}$$

where $\eta \geq \epsilon_m$ is the relative error in calculating $f(x)$ and ϵ_m is the machine epsilon (which is the smallest positive number such that $1 + \epsilon_m > 1$ on the computer in question), and for the central differences

$$h_i = \eta^{1/3} x_i \tag{A.20}$$

see (Dennis & Schnabel, 1983). The optimal step size is of course the balance between minimal truncation error and rounding errors. The truncation error vanishes as $h_i \to 0$ but because of finite precision arithmetic the rounding error then will grow.

Usually the forward difference will provide approximate gradients of acceptable accuracy unless $||g(x)||$ is small. Since the gradient approaches zero at the solution of an unconstrained problem, this means that the gradient approximation should switch to the (more expensive) central difference when coming closer to the optimum. The switch should be effectuated when the line search (see Section A.2.3) fails in finding a lower point. This tends to happen only when the relative error in the overall gradient vector, based on finite differences, becomes too large (Gill, Murray, Saunders, & Wright, 1983).

A.2.2 BFGS-update

The most effective secant approximation $\mathbf{B}_n$ to the Hessian in (A.16) is the BFGS method (Dennis & Schnabel, 1983)

$$\mathbf{B}_{n+1} = \mathbf{B}_n + \frac{\mathbf{y}_n \mathbf{y}_n^T}{\mathbf{y}_n^T \mathbf{s}_n} - \frac{\mathbf{B}_n \mathbf{s}_n \mathbf{s}_n^T \mathbf{B}_n}{\mathbf{s}_n^T \mathbf{B}_n \mathbf{s}_n} \tag{A.21}$$

where $\mathbf{y}_n = \mathbf{g}(\mathbf{x}_{n+1}) - \mathbf{g}(\mathbf{x}_n)$ and $\mathbf{s}_n = \mathbf{x}_{n+1} - \mathbf{x}_n$. The necessary and sufficient conditions for this formula to have a positive definite solution $\mathbf{B}_{n+1}$ is that $\mathbf{B}_n$ is positive definite and

$$\mathbf{y}_n^T \mathbf{s}_n > 0 \tag{A.22}$$

The soft line search will meet this last demand. Since the Hessian is symmetric and positive definite we prefer to write it in terms of its Cholesky factors (square root free)

$$\mathbf{B}_n = \mathbf{L}_n \mathbf{D}_n \mathbf{L}_n^T \tag{A.23}$$

with $\mathbf{L}_n$ lower triangular and $\mathbf{D}_n$ diagonal. Instead of actually calculate and factor (A.21) at each iteration, the Cholesky factorization of $\mathbf{B}_n$ can be updated using the $\mathbf{QR}$ scheme mentioned in Section A.1.2.

A.2.3 Soft Line Search

The idea of the soft line search is to choose a $\lambda_n > 0$ for

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \lambda_n \mathbf{s}_n \tag{A.24}$$

to assure that the next iterate decreases $f(\mathbf{x})$ sufficiently and that the condition (A.22) is satisfied. $\mathbf{s}_n$ is the secant direction obtained from (A.16) with the BFGS update of the Hessian, this step will always be a descent direction: $\mathbf{g}(\mathbf{x}_n)^T \mathbf{s}_n < 0$. Often $\lambda_n = 1$ will satisfy the demands, so that the line search method reduces to the secant method (A.16).

It has been shown that the line search will be globally convergent if each step satisfies two simple conditions. The first is that the decrease in $f(\mathbf{x})$ is sufficient in relation to the length of the step $\mathbf{s}'_n = \lambda_n \mathbf{s}_n$; the relation

$$f(\mathbf{x}_{n+1}) < f(\mathbf{x}_n) + \alpha \mathbf{g}(\mathbf{x}_n)^T \mathbf{s}'_n \tag{A.25}$$

where $\alpha \in]0, 1[$ is some constant, is chosen to implement this condition. The second condition is that the step is not too short. The equation

$$\mathbf{g}(\mathbf{x}_{n+1})^T \mathbf{s}'_n \geq \beta \mathbf{g}(\mathbf{x}_n)^T \mathbf{s}'_n \tag{A.26}$$

where $\beta \in]\alpha; 1[$ will implement this condition. This last expression and $\mathbf{g}(\mathbf{x}_n)^T \mathbf{s}'_n < 0$ imply

$$\mathbf{y}_n^T \mathbf{s}'_n = (\mathbf{g}(\mathbf{x}_{n+1}) - \mathbf{g}(\mathbf{x}_n))^T \mathbf{s}'_n \geq (\beta - 1)\mathbf{g}(\mathbf{x}_n)^T \mathbf{s}'_n > 0 \tag{A.27}$$

which is the condition (A.22).

A.3 Kalman Filter

For linear models the Kalman filter provides the exact solution for the filtering problem discussed in Chapter 4, see e.g. (Åström, 1970). The equations for the Kalman filter are given in Section 4.1.4 but is repeated below. The routine calculates the conditional mean $\hat{\mathbf{y}}_{k|k-1}$ and the conditional variance $\mathbf{R}_{k|k-1}$ to be used for the estimation. The Kalman filter performs the optimal (minimum variance) linear updating and prediction of the state variables. If the model is time invariant the model can be discretized before filtering. If the time dependency is weak compared to the dominating eigenvalues of the system, this implementation of the Kalman filter may also be used for time varying systems, by discretizing the continuous model at each sampling instant, assuming that $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{G}$ are constant within the sampling interval.

For the discrete time model specified in Section 4.1.4 the equations for *updating* the estimate of the state $\mathbf{x}$ becomes

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}) \tag{A.28}$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k \mathbf{R}_{k|k-1} \mathbf{K}_k^T \tag{A.29}$$

where $\mathbf{K}_k$ is given by

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{C}^T\mathbf{R}_{k|k-1}^{-1} \tag{A.30}$$

The formulas for *prediction* becomes

$$\hat{\mathbf{x}}_{k+1|k} = \mathbf{\Phi}\hat{\mathbf{x}}_{k|k} + \mathbf{\Gamma}\mathbf{u}_k \tag{A.31}$$

$$\hat{\mathbf{y}}_{k+1|k} = \mathbf{C}\hat{\mathbf{x}}_{k+1|k} + \mathbf{D}\mathbf{u}_{k+1} \tag{A.32}$$

$$\mathbf{P}_{k+1|k} = \mathbf{\Phi}\mathbf{P}_{k|k}\mathbf{\Phi}^T + \mathbf{\Lambda} \tag{A.33}$$

$$\mathbf{R}_{k+1|k} = \mathbf{C}\mathbf{P}_{k+1|k}\mathbf{C}^T + \mathbf{S} \tag{A.34}$$

The formulas require some initial values, which describes the prior knowledge about the states of the system in terms of the prior mean and variance $\hat{\mathbf{x}}_{1|0} = \mu_0$ and $\mathbf{P}_{1|0} = \mathbf{V}_0$. The matrix $\mathbf{P}_{k+1|k}$ is the variance of the one-step prediction of the state, $\mathbf{x}$, of the system.

It is well known that the Kalman filter in some situations is numerically unstable. The problems arise when some of the variances, because of rounding errors, become non-positive definite. Therefore careful handling of the equations for the variances (A.29), (A.30), (A.33) and (A.34) is needed in order to numerically stabilize the Kalman filter. Since all variances should be symmetric and positive definite, it is desirable to use their Cholesky factorization. We use the $\mathbf{LDL}^T$-factorization, i.e. the square root free Cholesky decomposition, where $\mathbf{L}$ is unit lower matrix and $\mathbf{D}$ is diagonal.

An equation for updating a factorized matrix is

$$\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{G}\mathbf{D}_g\mathbf{G}^T \tag{A.35}$$

where $\tilde{\mathbf{A}}$ is known from other considerations to be positive definite, and $\mathbf{D}_g$ is a diagonal matrix. Thus, it is necessary to compute a unit lower triangular matrix $\tilde{\mathbf{L}}$ and a diagonal matrix $\tilde{\mathbf{D}}$ with $\tilde{d}_i > 0$ such that

$$\tilde{\mathbf{A}} = \tilde{\mathbf{L}}\tilde{\mathbf{D}}\tilde{\mathbf{L}}^T \tag{A.36}$$

Algorithms to solve this problem are outlined in Section A.1.2. It is obvious that equation (A.33) and (A.34) easily are brought into this form. Equation (A.29) can be rewritten as

$$\begin{aligned}
\mathbf{P}_{k|k} &= \mathbf{P}_{k|k-1} - \mathbf{K}_k \mathbf{R}_{k|k-1} \mathbf{K}_k^T \quad \Leftrightarrow \\
\mathbf{P}_{k|k} &= \mathbf{P}_{k|k-1} - \mathbf{P}_{k|k-1} \mathbf{C}^T \mathbf{R}_{k|k-1}^{-1} \mathbf{C} \mathbf{P}_{k|k-1}^T \quad \Leftrightarrow \\
\tilde{\mathbf{L}}\tilde{\mathbf{D}}\tilde{\mathbf{L}}^T &= \mathbf{L}\mathbf{D}\mathbf{L}^T - \mathbf{L}\mathbf{D}\mathbf{L}^T \mathbf{C}^T (\mathbf{L}_r \mathbf{D}_r \mathbf{L}_r^T)^{-1} \mathbf{C} (\mathbf{L}\mathbf{D}\mathbf{L}^T)^T \quad \Leftrightarrow \\
\tilde{\mathbf{L}}\tilde{\mathbf{D}}\tilde{\mathbf{L}}^T &= \mathbf{L}(\mathbf{D} - \mathbf{G}\mathbf{D}_r^{-1}\mathbf{G}^T)\mathbf{L}^T
\end{aligned} \tag{A.37}$$

where

$$\mathbf{G} = \mathbf{D}\mathbf{L}^T\mathbf{C}^T\mathbf{L}_r^{-T} \tag{A.38}$$

The expression $(\mathbf{D} - \mathbf{G}\mathbf{D}_r^{-1}\mathbf{G}^T)$ in (A.37) are in the form (A.35), and can thus be solved for the factors $\bar{\mathbf{L}}\bar{\mathbf{D}}\bar{\mathbf{L}}^T$, and we have $\tilde{\mathbf{L}} = \mathbf{L}\bar{\mathbf{L}}$ and $\tilde{\mathbf{D}} = \bar{\mathbf{D}}$.

This implementation of the Kalman filter is able to handle the multiple-input, multiple-output case with a high degree of accuracy and stability, see e.g. (Bierman, 1977).

A.4 Extended Kalman Filter

Let the model be described by the stochastic differential equation

$$d\mathbf{x}_t = \mathbf{f}(\mathbf{x}_t, \mathbf{u}_t, \theta, t)dt + \mathbf{G}(\theta, t)d\beta_t \tag{A.39}$$

with β being a standard Wiener process. The observations $\mathbf{y}_k$ are taken at discrete time instants, t_k

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k, \mathbf{u}_k, \theta, t_k) + \mathbf{e}_k \tag{A.40}$$

where $\mathbf{e}$ is a Gaussian white noise process independent of β, and $\mathbf{e}_k \sim N(0, \mathbf{S}(\theta, t_k))$. In this case the extended Kalman filter is used as a first order approximative filter. Being linearized about $\hat{\mathbf{x}}_t$ the state and covariance propagation equations have a structure similar to the Kalman filter propagation equations for linear systems. Hence, we are able to reuse the numerical stable routines implemented for the Kalman filter from the previous section.

The necessary modifications of the equations in the previous section are the following. The matrix $\mathbf{C}$ is the linearization of the measurement equation,

$$\mathbf{C}(\hat{\mathbf{x}}_{k|k-1}, \mathbf{u}_k, \theta, t_k) = \frac{\partial \mathbf{h}}{\partial \mathbf{x}}\Big|_{x=\hat{x}_{k|k-1}}, \tag{A.41}$$

and $\mathbf{A}$ is the linearization of the system equation,

$$\mathbf{A}(\hat{\mathbf{x}}_t, \mathbf{u}_t, \theta, t) = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\Big|_{x=\hat{x}_t}, \tag{A.42}$$

and $\mathbf{\Phi}$ is the discrete system matrix calculated as a transformation of $\mathbf{A}$, see the following section. The prediction of the output, Equation A.32, is replaced by

$$\hat{\mathbf{y}}_{k+1|k} = \mathbf{h}(\hat{\mathbf{x}}_{k+1|k}, \mathbf{u}_{k+1}, \theta, t_{k+1}) \tag{A.43}$$

The formulas for prediction of mean and covariance of the state-vector are normally given by,

$$d\hat{\mathbf{x}}_{t|k}/dt = \mathbf{f}(\hat{\mathbf{x}}_{t|k}, \mathbf{u}_t, \theta, t))\,, \qquad t \in [t_k, t_{k+1}[\tag{A.44}$$

$$d\mathbf{P}_{t|k}/dt = \mathbf{A}(\hat{\mathbf{x}}_{t|k}, \mathbf{u}_t, \theta, t)\,\mathbf{P}_{t|k} + \mathbf{P}_{t|k}\mathbf{A}^T(\hat{\mathbf{x}}_{t|k}, \mathbf{u}_t, \theta, t) + \mathbf{G}(\theta, t)\mathbf{G}^T(\theta, t)\,, \qquad t \in [t_k, t_{k+1}[\tag{A.45}$$

where $\mathbf{A}$ is given by (A.42). In order to make the integration of (A.44) and (A.45) computational feasible and numerically stable for stiff systems, the time interval $[t_k, t_{k+1}[$ is sub sampled and the equations are linearized about the state estimate at the given sub sampling time. For the state propagation the equation becomes

$$d\hat{\mathbf{x}}_t/dt = \mathbf{f}(\hat{\mathbf{x}}_j) + \mathbf{A}(\hat{\mathbf{x}}_j)\{\hat{\mathbf{x}}_t - \hat{\mathbf{x}}_j\}\,, \qquad t \in [t_j, t_{j+1}[\tag{A.46}$$

$$= \mathbf{A}(\hat{\mathbf{x}}_j)\,\hat{\mathbf{x}}_t + \{\mathbf{f}(\hat{\mathbf{x}}_j) - \mathbf{A}(\hat{\mathbf{x}}_j)\,\hat{\mathbf{x}}_j\}\,, \; t \in [t_j, t_{j+1}[\tag{A.47}$$

where $[t_j, t_{j+1}[$ is one of the sub intervals of the sampling interval $[t_k, t_{k+1}[$, assuming the sampling interval has been divided in n_s sub intervals. In these derivations only the state dependency is mentioned for clarity. Equation A.47 is a linear ordinary differential equation which has the exact solution

$$\hat{\mathbf{x}}_{j+1} = \hat{\mathbf{x}}_j + (e^{\mathbf{A}(\hat{\mathbf{x}}_j)\tau_s} - \mathbf{I})(\mathbf{A}(\hat{\mathbf{x}}_j)^{-1}\mathbf{f}(\hat{\mathbf{x}}_j)) \tag{A.48}$$

$$= \hat{\mathbf{x}}_j + (\mathbf{\Phi}_s(\hat{\mathbf{x}}_j) - \mathbf{I})(\mathbf{A}(\hat{\mathbf{x}}_j)^{-1}\mathbf{f}(\hat{\mathbf{x}}_j)) \tag{A.49}$$

where $\tau_s = t_{j+1} - t_j = \tau/n_s$, and τ is the sampling time. Efficient algorithms to calculate the matrix exponential in (A.48) is described in the next section. Correspondingly the equation for the state covariance becomes

$$\mathbf{P}_{j+1} = \mathbf{\Phi}_s(\hat{\mathbf{x}}_j)\mathbf{P}_j\mathbf{\Phi}_s(\hat{\mathbf{x}}_j)^T + \mathbf{\Lambda}_s(\hat{\mathbf{x}}_j)\,, \tag{A.50}$$

which is similar to (A.33). The algorithm to solve (A.44) and (A.45) is use $\hat{\mathbf{x}}_{k|k}$ and $\hat{\mathbf{P}}_{k|k}$ as starting values for (A.49) and (A.50) and then perform n_s iterations of (A.49) and (A.50) simultaneously. This algorithm has the advantage of being numerically stable for stiff systems and still computationally efficient, since the fast and stable routines of the linear Kalman filter can be used.

A.5 Discretizing the Model

When using the Kalman filter from the previous sections we need to discretize the model either once for a given set of parameters, for time invariant models or at each sampling instant, or more frequently, for time invariant models and for the extended Kalman filter. Hence we need to calculate the following quantities in an efficient way

$$\begin{aligned}&\mathbf{\Phi}(\tau) = e^{\mathbf{A}\tau} \ ; \qquad \Gamma(\tau) = \int_0^\tau e^{\mathbf{A}s}\mathbf{B}\,ds \ ; \\ &\mathbf{\Lambda}(\tau) = \int_0^\tau \mathbf{\Phi}(s)\mathbf{G}\mathbf{G}^T\mathbf{\Phi}(s)^T\,ds\end{aligned} \tag{A.51}$$

where τ is the sampling time, i.e. we need to calculate the exponential of a matrix and integrals involving the matrix exponential.

A.5.1 Matrix Exponential

An approach, which has proven to be very successful for calculating $\exp(\mathbf{A})$ uses diagonal Padé approximation with repeated squaring, see (Moler & van Loan, 1978). The estimation of the matrix exponential is given by

$$e^{\mathbf{A}} = \left(\mathbf{R}_{qq}(\mathbf{A}/2^j)\right)^{2^j} \ , \qquad q, j \geq 0 \tag{A.52}$$

where $\mathbf{R}_{qq}(z)$ is the (q,q) Padé approximation to $\exp(z)$

$$\mathbf{R}_{qq}(z) = \frac{\sum_{k=0}^{q} c_k z^k}{\sum_{k=0}^{q} c_k(-z)^k} \ , \qquad c_k = \frac{(2q-k)!q!}{(2q)!k!(q-k)!} \tag{A.53}$$

The scaling by 2^j followed by the repeated squaring greatly enhances the numerical properties of ordinary Padé approximation (Moler & van Loan, 1978). Usually j is chosen as the smallest nonnegative integer such that $||\mathbf{A}||/2^j \leq 1/2$, where $||\cdot||$ denotes the Frobenius norm

$$||\mathbf{A}|| = \Big(\sum_i \sum_j |A_{ij}|^2\Big)^{1/2} \tag{A.54}$$

Other norms are possible, but the Frobenius norm is convenient for both practical and theoretical reasons. According to Moler & van Loan (1978) the resulting algorithm is one of the most computationally efficient.

A.5.2 Integrals Involving Matrix Exponentials

The most effective way to calculate the integrals involved in (A.51) is by computing the exponential of a certain block triangular matrix and combining the resulting sub matrices. The principle and algorithms for the calculation are given in (van Loan, 1978), where the method is also compared with alternative techniques, e.g. Simpson integration.

By calculating the exponential of the following block triangular matrix

$$\exp\left(\begin{bmatrix} -\mathbf{A} & \mathbf{R} & 0 \\ 0 & \mathbf{A}^T & \mathbf{I} \\ 0 & 0 & 0 \end{bmatrix} \tau\right) = \begin{bmatrix} \mathbf{F}_1(\tau) & \mathbf{G}_1(\tau) & \mathbf{H}_1(\tau) \\ 0 & \mathbf{F}_2(\tau) & \mathbf{G}_2(\tau) \\ 0 & 0 & 0 \end{bmatrix} \tag{A.55}$$

where $\mathbf{R} = \mathbf{G}\mathbf{G}^T$ from (A.51), we are able to calculate the quantities of (A.51) by simple algebraic operations. It can be shown that

$$\mathbf{\Phi}(\tau) = \mathbf{F}_2(\tau)^T \tag{A.56}$$

$$\mathbf{\Gamma}(\tau) = \mathbf{G}_2(\tau)^T\mathbf{B} \tag{A.57}$$

$$\mathbf{\Lambda}(\tau) = \mathbf{F}_2(\tau)^T\mathbf{G}_1(\tau) \tag{A.58}$$

see (van Loan, 1978). For calculation of the matrix exponential the Padé approximation described in the previous section is used. However, in the interest of efficiency, the algorithm does not repeatedly square the matrix $R_{qq}(\mathbf{A}/2^j)$ as suggested by (A.52). Instead doubling formulas for the terms (A.56), (A.57) and (A.58) are repeatedly exploited, see (van Loan, 1978).

The resulting algorithm for calculating the integrals of (A.51) is very efficient and superior to other algorithms concerning computational speed as well as accuracy according to van Loan (1978).

A.6 Random Number Generation

For applications of stochastic simulation sequences of pseudorandom numbers are needed. Usually they are generated by a deterministic algorithm which produces a sequence of numbers that behave as a realization of a sequence of independent identically distributed random variables having a uniform distribution on the unit interval. In a second step these standard pseudorandom numbers are transformed to fit some other distribution, e.g. the Gaussian distribution, (Devroye, 1986). It is very important to be aware of the algorithm calculating the random numbers, because they are the foundations of our simulations. “Often problems on higher levels are traced back to faulty foundations”, (Ripley, 1987).

Most frequently pseudorandom numbers are generated by the *linear congruential method*, which is defined by

$$x_i \equiv (a x_{i-1} + b) \bmod m \qquad \text{(A.59)}$$

for a multiplier a, a shift b, and modulus m, all integers. If $b = 0$ the generator is called multiplicative congruential. The pseudorandom sequence

$\{u_i\}$, taking values in $[0, 1]$, is determined by $u_i = x_i/m$ once the seed x_0 is chosen. For the purpose of obtaining as large a period length as possible, m should be chosen large. Since the generated sequence $\{u_i\}$ is periodical with period length $\leq m$, the parameters a and b should be chosen, so the generator achieves its maximum period length. The maximum period length for a *multiplicative congruential generator* is $m - 1$.

If we consider the k-tuples $(u_i, \cdots, u_{i+k-1})$ from the generator, as points in the hypercupe $[0, 1]^k$, then can be shown that the k-tuples from a maximal-period linear congruential generator will always lie on a finite number of hyperplanes in $[0, 1]^k$, see (Ripley, 1987, chapter 2). In order to obtain a generator which "behaves" sufficiently randomly the parameters a and b should be chosen, such that the sequence of points $(u_i, \cdots, u_{i+k-1})$ does not concentrate on too few hyperplanes, for reasonable dimensions $k = 2, 3, 4$ and preferably for $k \leq 10$, i.e. the k-tuples are as uniformly distributed as possible in $[0, 1]^k$ for these dimensions, recommended by Ripley (1987, p. 45). Various measures of the "granularity" of the lattice generated by the hyperplanes, have been developed. One method is to measure the maximum distance between the hyperplanes. This distance should be as little as possible for a good generator. Another way of analyzing the lattice structure is by calculation of a basis of shortest vectors $e_1, e_2, \cdots, e_k$. The unit cell of the lattice is defined as e_1 being the shortest nonzero vector in the lattice, e_2 as the shortest vector independent of e_1, e_3 as the shortest vector independent of e_1 and e_2, and so on. Hence we have that $||e_1|| \leq ||e_2|| \leq \cdots \leq ||e_k||$, and the Beyer ratio $q_k = ||e_1||/||e_k||$ measures the "uniformity" of the lattice. One is interested in generators with q_k close to one, for reasonable dimensions, k. Procedures to calculate these lattice constants for congruential generators may be found in (Ripley, 1987).

As a generalization of the linear congruential method Knuth (1981) introduced multi-recursive generators

$$x_i \equiv (a_1 x_{i-1} + \cdots + a_r x_{i-r}) \bmod m \tag{A.60}$$

for recursion depth r and modulus m. The maximum period length of this generator is $m^r - 1$. If we confine the modulus in (A.60) to $m = 2$, it is very easy to implement by the use of a shift register, hence the algorithm is called a *shift-register generator*.

Both the linear congruential method of (A.59) and the generalized method (A.60) suffer from the lattice structure described above. In an attempt to get rid of the lattice structure of the generators, the *inversive congruential method* have been introduced, see (Lehn, 1991). It is defined by

$$x_i \equiv (a\bar{x}_{i-1} + b) \bmod m \tag{A.61}$$

where m is a prime modulus, the starting value x_0 as well as the coefficients a and b are integers between 0 and $m-1$. $\bar{x}$ is defined in the following way: for any $x \in \{1, 2, \cdots, m-1\}$ there is exactly one $\bar{x} \in \{1, 2, \cdots, m-1\}$ where $x\bar{x} = 1 \bmod m$. Together with $\bar{0} = 0$ this defines $\bar{x}$. In (Lehn, 1991) is given an algorithm for calculating $\bar{x}$, witch asymptotically need $O(\log m)$ steps of iteration. Therefore, the generation of inversive congruential pseudorandom numbers takes more time than is needed with the linear congruential method. The inversive congruential method do not create a lattice structure as the linear congruential methods, and have in this sense a behavior closer to truly random numbers, see (Lehn, 1991).

A number of statistical tests exists, which can be applied to a sequence of the output $(u_1, \cdots, u_n)$ from a pseudorandom generator. The relevant tests for random number generators include tests for independence, e.g.

gaps test, runs test and permutation tests, see (Ripley, 1987), others may be found in (Brockwell & Davis, 1987; Kendall & Stuart, 1979), and tests for uniformity, e.g. Kolmogorov-Smirnov test for distribution, see (Kendall & Stuart, 1979). The theoretical tests (Beyer ratios etc.) discussed earlier have been found to be more powerful than statistical tests in the sense that "good" generators by the theoretical criteria have been found to fail the statistical tests no more often than would be expected by chance. Nevertheless it is always worth conducting some empirical tests to check that the generator has been implemented correctly, following Ripley (1987, pp. 43–45).

Different approaches of transforming the standard pseudorandom numbers to some given non-uniform distribution is discussed in (Devroye, 1986). From the pair (u_1, u_2) of independent random numbers, uniformly distributed on $[0, 1]$ we have used the Box and Muller transformation:

$$(\theta\,,\, r) = (2\pi u_1\,,\, \sqrt{-2\log u_2}) \quad \text{and} \tag{A.62}$$

$$(y_1\,,\, y_2) = (r\cos\theta\,,\, r\sin\theta) \tag{A.63}$$

for generating independent standard normal deviates (y_1, y_2).

A.7 Validation of The Estimation Tool - Predator/Prey Relations

In order to validate the implementation of the estimation tool and specifically the extended Kalman filter, an example using a nonlinear model is considered. The example is a classical type of nonlinear system describing the temporal oscillations of a population of predators and their prey in a localized geographic region. The prey population is denoted by $N_1(t)$ and the predator population by $N_2(t)$. It is assumed that the growth rate of prey, in the absence of predators, is $aN_1(t)$, while the predator multiplication rate is $cN_1(t)N_2(t)$. Further, is is assumed that the loss rate of prey is proportional to the numbers of prey and predators, i.e. loss rate of prey equals $bN_1(t)N_2(t)$, while the loss rate of predators equals their death rate $dN_2(t)$. Putting together these assumptions, the dynamics of the predator - prey interaction are described by the deterministic Lotka-Volterra equations

$$\dot{N}_1(t) = aN_1(t) - bN_1(t)N_2(t) , \tag{A.64}$$

$$\dot{N}_2(t) = cN_1(t)N_2(t) - dN_2(t) , \tag{A.65}$$

where a, b, c and d all are positive constants. The point $N_1^* = d/c$, $N_2^* = a/b$ is the only nontrivial equilibrium of the Lotka-Volterra system. If the initial condition $(N_1(0), N_2(0)) \neq (N_1^*, N_2^*)$, the trajectory of the system is a closed orbit as depicted in Figure A.1. In the figure, the trajectory is moving anti-clockwise. Thus, for any given initial condition, the populations of predator and prey will oscillate cyclically. Neither species will die out, nor will it grow indefinitely. Furthermore, except for the improbable initial state (N_1^*, N_2^*), the populations will not remain constant.

The deterministic Lotka-Volterra model is modified in order to give a more likely picture of a real world system. It is assumed that the rate of change of each population is influenced by an additive random term representing the effect of other factors not included in the model, such as other predators in the food chain and weather variables. This results in a stochastic Lotka-Volterra model

$$dN_1(t) = (aN_1(t) - bN_1(t)N_2(t))\,dt + \sigma_1\,d\beta_1(t) , \tag{A.66}$$

$$dN_2(t) = (cN_1(t)N_2(t) - dN_2(t))\,dt + \sigma_2 d\beta_2(t) , \tag{A.67}$$

where $\beta_1(t)$ and $\beta_2(t)$ are mutually independent standard Wiener pro-

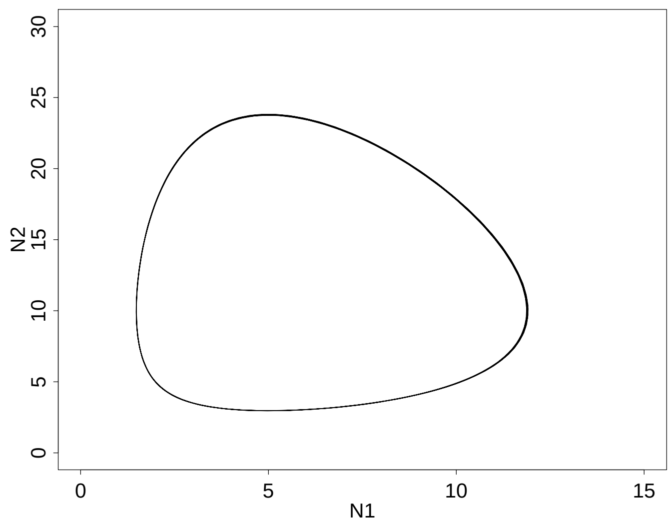


Figure A.1. *Phase-plane of the deterministic Lotka-Volterra system.*

cesses and σ_1 and σ_2 are positive constants giving the levels of the noise. If we are able to observe the populations, by measuring the system at given time instants, the measurements are assumed to by influenced by a random error, i.e. the measurement equation is

$$N_{1,m}(t_k) = N_1(t_k) + \sigma_{1,m} e_1(t_k) \ , \tag{A.68}$$

$$N_{2,m}(t_k) = N_2(t_k) + \sigma_{2,m} e_2(t_k) \ , \tag{A.69}$$

where $e_1(t_k)$ and $e_2(t_k)$ are standard Gaussian white noise and $\sigma_{1,m}$ and $\sigma_{2,m}$ are positive constants giving the levels of the noise. An example of the measurements of such a system is simulated, using stochastic simulation, and the timeseries of measured populations are shown in Figure A.2. In this example the following dimensionless variables were chosen, $a = 10$, $b = 1$, $c = 2$, $d = 10$, $\sigma_1^2 = \sigma_2^2 = 0.3$, $\sigma_{1,m}^2 = \sigma_{2,m}^2 = 0.03$, and the sampling time, $T_s = 0.005$, and initial populations, $(N_1(0), N_2(0)) = (2, 5)$. A phase-plane

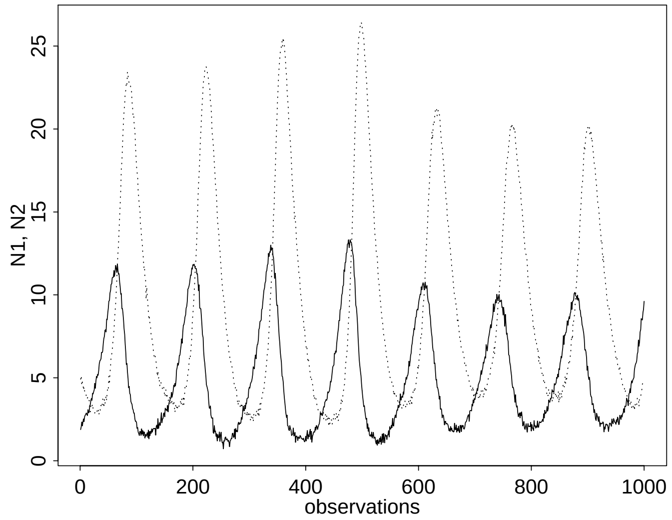


Figure A.2. *Timeseries plot of the populations* N1 *and* N2 *for the stochastic Lotka-Volterra system.*

of the same realization as in Figure A.2 is shown in Figurelv-stoc2. The qualitative structure of the trajectory is the same as for the deterministic system and it is still a stable oscillating system.

For the purpose of validating the estimation software, in this case is CTLSM (Melgaard & Madsen, 1993), and specifically the extended Kalman filter implementation, 50 sequences of stochastic independent realizations of the system was simulated. Each sequence with the length of 500 observations of the populations. In Table A.1 the results from estimation of the parameters of the model from the 50 series are summarized. For all parameters, the mean of the estimated values are given, which can be compared to the simulated values. Also the empirical variance and the mean of the estimated variances of the parameters are given. A comparison of these values

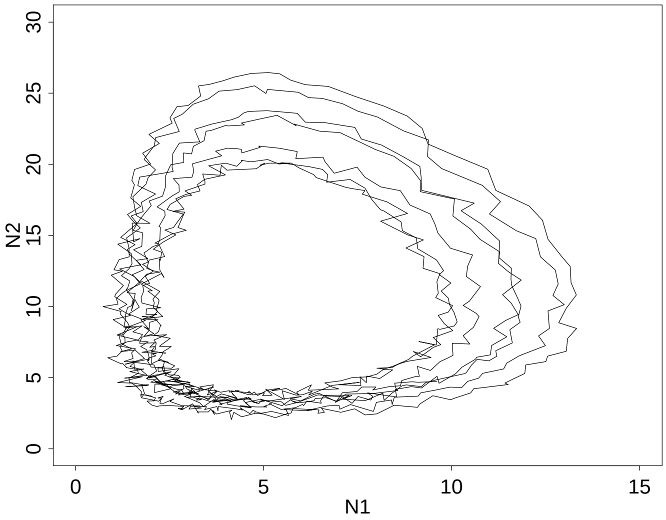


Figure A.3. *Phase-plane of the stochastic Lotka-Volterra system.*

will indicate if the tool is able to estimate the right uncertainty of the parameters. The columns of the table are: x_{sim}, the simulated values, $\bar{x}$, the mean of the estimated values, s_x^2 is the empirical variance of the estimated parameters, $\bar{s^2}$ is the mean of the estimated variance of the parameters, F-stat. is an F statistic given by $Z_F = s_x^2/\bar{s^2}$ and $|t|$-stat. is a t statistic given by $Z_t = |\bar{x} - x_{sim}|/(s_x\sqrt{n_e})$.

In order to verify if the variance of the parameters provided by the estimation tool is equal to the empirical variance, one wish to test the hypotheses

$$\begin{aligned} H_0 : &\quad s_x^2 = \bar{s^2} \\ H_1 : &\quad s_x^2 \neq \bar{s^2} \,. \end{aligned}$$

Under H_0 we have in this case that $Z_F \sim F(49, \infty)$. The critical set for this test is $\{z < F(49,\infty)_{\alpha/2} \vee z > F(49,\infty)_{1-\alpha/2}\}$ on level α. By choosing

Parameter	x_{sim}	$\bar{x}$	s_x^2	$\bar{s^2}$	F-stat.	|t|-stat.
a	10.000	9.985	8.568e-3	9.589e-3	0.8936	1.151
b	1.0000	0.9990	6.453e-5	8.267e-5	0.7805	0.914
c	2.0000	2.0012	1.253e-4	1.095e-4	1.1441	0.758
d	10.000	10.001	4.756e-3	4.481e-3	1.0614	1.446
$N_1(0)$	2.0000	2.0066	5.426e-3	5.888e-3	0.9215	0.630
$N_2(0)$	5.0000	5.0175	7.893e-3	5.902e-3	1.3373	1.393
σ_1^2	0.3000	0.2909	2.883e-3	3.529e-3	0.8169	1.196
σ_2^2	0.3000	0.2896	9.011e-3	7.981e-3	1.1290	0.773
$\sigma_{1,m}^2$	0.03000	0.02966	3.229e-6	3.855e-6	0.8377	1.346
$\sigma_{2,m}^2$	0.03000	0.03047	4.593e-6	4.184e-6	1.0978	1.534

Table A.1. *Results from estimation of the* $n_e = 50$ *series from the stochastic Lotka-Volterra system.*

$\alpha = 0.1$ we obtain the critical set $\{z < 0.74 \vee z > 1.35\}$. It is seen from Table A.1 that we cannot reject H_0 for any parameter on the chosen level.

Another test is performed in order to verify that the estimated parameters are un-biased. The following hypotheses are tested

$$\begin{aligned} H_0 &: \quad \bar{x} = x_{sim} \\ H_1 &: \quad \bar{x} \neq x_{sim} \,. \end{aligned}$$

Under H_0 the distribution of the test statistic is $Z_t \sim t(49)$. The critical set for this test is $\{z > t(49)_{1-\alpha/2}\}$ on level α. For $\alpha = 0.1$, the critical set is $\{z > 2.0\}$, thus from Table A.1 we cannot reject H_0 for any of the parameters on the chosen level.

References

Akaike, H. (1976). Canonical correlation analysis of time series and the use of an information criterion. In Mehra, R. K. & Lainiotis, D. G. (Eds.), *System Identification: Advances and Case Studies*, pp. 27–96. Academic Press, New York.

Aquino, C. F. (1981). Transient A/F control characteristics of the 5 liter central fuel injection engine.. SAE Technical Paper No 810494.

Åström, K. J. (1970). *Introduction to Stochastic Control Theory.* Academic Press, New York.

Bierman, G. J. (1977). *Factorization Methods for Discrete Sequential Estimation.* Academic Press, New York.

Björck, A. (1967). Solving linear least squares problems by Gram-Schmidt orthogonalization. *BIT*, **7**, 1–21.

Bohlin, T. (1978). Maximum-power validation of models without higher-order fitting. *Automatica*, **14**, 137–146.

Bohlin, T. (1984). Computer-aided grey-box validation. Tech. rep. TRITA-REG-8403, Department of Automatic Control, Royal Institute of Technology, Stockholm, Sweden.

Bohlin, T. (1989). The fundamentals of modelling and identification. Tech. rep. TRITA-REG-89/00002, Department of Automatic Control, Royal Institute of Technology, Stockholm, Sweden.

Bohlin, T. (1994). A case study of grey box identification. *Automatica*, **30**(2), 307–318.

Box, G. E. P. & Jenkins, J. M. (1976). *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco.

Brockwell, P. J. & Davis, R. A. (1987). *Time Series: Theory and Methods*. Springer Series in Statistics. Springer-Verlag, New York.

Collacott, R. A. (1977). *Mechanical Fault Diagnosis and Condition Monitoring*. Chapman and Hall.

Dennis, J. E. & Schnabel, R. B. (1983). *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Prentice-Hall, Englewood Cliffs, New Jersey.

Devroye, L. (1986). *Non-Uniform Random Variate Generation*. Springer-Verlag, Berlin.

Doob, J. L. (1953). *Stochastic processes*. John Wiley, New York.

Fedorov, V. V. (1972). *Theory of Optimal Experiments*. Academic Press, New York.

Fletcher, R. & Powell, J. D. (1974). On the modification of $\mathrm{LDL}^{\mathrm{t}}$ factorizations. *Math. Comp.*, **28**, 1067–1087.

Gard, T. C. (1988). *Introduction to Stochastic Differential Equations*. Marcel Dekker, New York.

Gelb, A. (1974). *Applied Optimal Estimation*. MIT Press, New York.

Gevers, M. & Ljung, L. (1986). Optimal experiment designs with respect to the intended model application. *Automatica*, **22**(5), 543–554.

Gihman, I. I. & Skorohod, A. V. (1974,1975,1979). *The theory of stochastic processes*, Vol. 1,2,3. Springer, Berlin.

Gill, P. E., Murray, W., Saunders, M. A., & Wright, M. H. (1983). Computing forward difference intervals for numerical optimization. *SIAM J. Sci. Stat. Comput.*, **4**, 310–321.

Godfrey, K. R. (1980). Correlation methods. *Automatica*, **16**, 527–534.

Golub, G. H. & van Loan, C. F. (1983). *Matrix Computations*. North Oxford Academic, Oxford.

Goodson, R. E. (1970). Distributed system simulation using infinite product expansions.. *Simulation*, December.

Goodwin, G. C. & Payne, R. L. (1977). *Dynamic System Identification: Experiment Design and Data Analysis*. Academic Press, New York.

Graebe, S. (1989). On gray box identification. Tech. rep. TRITA-REG 89/00004, Department of Automatic Control, Royal Institute of Technology, Stockholm, Sweden.

Graebe, S. (1990a). *Theory and Implementation of Gray Box Identification*. Ph.D. thesis, Department of Automatic Control, Royal Institute of Technology, Stockholm, Sweden.

Graebe, S. (1990b). Idkit, a software for gray box identification, mathematical reference. Tech. rep. TRITA-REG 90/00003, Department of Automatic Control, Royal Institute of Technology, Stockholm, Sweden.

Hammersten, S., van Hattem, D., Bloem, H., & Colombo, R. (1988). Passive solar component testing with identification methods. *Solar Energy*, **41**(1), 5–13.

Hendricks, E. & Sorenson, S. C. (1990). Mean value modelling of spark ignition engines. In *1990 American Control Conference*, Vol. 2, pp. 1882–1887, San Diego, California, USA.

Hendricks, E. (1992). *Identification and Estimation of Nonlinear Systems Using Physical Modelling*. Ph.D. thesis, IMSOR, Technical University of Denmark. No. 61.

Heywood, J. B. (1989). *Internal Combustion Engine Fundamentals*. McGraw-Hill, New York.

Holst, J., Holst, U., Madsen, H., & Melgaard, H. (1992). Validation of grey box models. In *IFAC Symposium on Adaptive Control and Signal Processing 92*, pp. 407–414.

Jazwinski, A. H. (1970). *Stochastic Processes and Filtering Theory*. Academic Press, New York.

Jordan, T. L. (1968). Experiments on error growth associated with some linear least-squares procedures. *Math. Comp.*, **22**, 576–588.

Kabaila, P. V. (1981). Estimation based on one step ahead prediction versus estimation based on multi-step ahead prediction. *Stochastics*, **6**, 43–55.

Karatzas, I. & Shreve, S. E. (1988). *Brownian Motion and Stochastic Calculus*. Springer-Verlag, New York.

Kay, I. E. (1978). Manifold fuel film effects in an si engine.. SAE Technical Paper No 780944.

Kendall, M. G. & Stuart, A. (1979). *The Advanced Theory of Statistics* (4th edition)., Vol. 2. Griffin, London.

Kloeden, P. E. & Platen, E. (1989). A survey of numerical methods for stochastic differential equations. *Stochastic Hydrol. Hydraul.*, **3**, 155 – 178.

Kloeden, P. E. & Platen, E. (1992). *Numerical Solution of Stochastic Differential Equations.* Springer-Verlag, Berlin.

Knuth, D. E. (1981). *The Art of Computer Programming. Volume 2: Seminumerical Algorithms.* (2nd edition). Addison-Wesley, Reading.

Kushner, H. (1967). *Stochastic stability and control.* Academic Press, New York.

Lawson, C. L. & Hanson, R. J. (1974). *Solving Least Squares Problems.* Prentice Hall, New Jersey.

Lehn, J. (1991). Pseudorandom number generators. Darmstadt, Fachbereich Mathematik, Preprint No. 1428.

Lindfors, A., Christoffersson, A., Roberts, R., & Anderlind, G. (1992). Estimation of the thermal properties of an insulating slab: a frequency domain approach. In *Workshop on Parameter Identification Methods and Physical Reality*, pp. 141–173. Institute for Systems Engineering and Informatics, EUR 14863 EN.

Ljung, G. & Box, G. (1978). On a measure of lack of fit in time series models. *Biometrika*, **65**, 297–303.

Ljung, L. & Caines, P. E. (1979). Asymptotic normality of prediction error estimators for approximate system models. *Stochastics*, **3**, 29–46.

Ljung, L. & Glad, T. (1994). On global identifiability for arbitrary model parametrizations. *Automatica*, **30**(2), 265–276.

Ljung, L. & van Overbeek, A. J. M. (1978). Validation of approximate models obtained from prediction error identification. In *Proc. 7th IFAC Congress, Helsinki*, pp. 1899–1906.

Ljung, L. (1978). Convergence analysis of parametric identification methods. *IEEE Transactions on Automatic Control*, **AC–23**(5), 770–783.

Ljung, L. (1979). Asymptotic behavior of the extended kalman filter as a parameter estimator for linear systems. *IEEE Transactions on Automatic Control*, **AC-24**(1), 36–50.

Ljung, L. (1987). *System Identification: Theory for the User*. Prentice-Hall, Englewood Cliffs, New Jersey.

van Loan, C. F. (1978). Computing integrals involving the matrix exponential. *IEEE Transactions on Automatic Control*, **AC–23**(3), 395–404.

Lütkepohl, H. (1985). Comparison of criteria for estimating the order of a vector autoregressive process. *J. Time Series Anal.*, **6**, 35–52.

Madsen, H. & Holst, J. (1993). Estimation of continuous time models for the heat dynamics of a building. *Energy and Buildings*, **Vol. ?** To be published.

Madsen, H. & Melgaard, H. (1991). The mathematical and numerical methods used in CTLSM - a program for ML-estimation in stochastic, continuous time dynamical models. Tech. rep. 7/1991, Institute of Mathematical Statistics and Operations Research, Technical University of Denmark, Lyngby, Denmark.

Madsen, H., Melgaard, H., & Holst, J. (1990). Identification of building performance parameters. In Bloem, J. (Ed.), *Workshop on Advanced Identification Tools in Solar Energy Research*, Non Nuclear Energy, pp. 37–60. Commission of the European Communities, DG XII.

Martin, R. D. & Yohai, V. J. (1985). Robustness in time series and estimating arma models. In Hannan, E. J., Krishnaiah, P. R., & Rao, M. M. (Eds.), *Handbook of Statistics*, Vol. 5, pp. 119–155. Elsevier Science Publishers, Amsterdam.

Maybeck, P. S. (1982). *Stochastic Models, Estimation, and Control*, Vol. 2. Academic Press, London.

Mehra, R. K. (1974). Optimal input signals for parameter estimation in dynamic systems - survey and new results. *IEEE Transactions on Automatic Control*, **AC-19**(6), 753–768.

Melgaard, H. & Madsen, H. (1993). CTLSM version 2.6 - a program for parameter estimation in stochastic differential equations. Tech. rep. 1/1993, Institute of Mathematical Statistics and Operations Research, Technical University of Denmark, Lyngby, Denmark.

Melgaard, H., Hendricks, E., & Madsen, H. (1990). Continuous identification of a four-stroke SI engine. In *1990 American Control Conference*, Vol. 2, pp. 1876–1881 San Diego, California, USA.

Melgaard, H., Madsen, H., & Holst, J. (1992a). Methods for parameter estimation in stochastic differential equations. In Platen, E. (Ed.), *Proceedings of the 1'st Workshop on Stochastic Numerics*, pp. 53–64. Institut für Angewandte Analysis und Stochastik, Berlin.

Melgaard, H., Madsen, H., & Holst, J. (1992b). Selection of optimal test sequences for identification of the heat dynamics of passive solar components. In *Thermal Performance of the Exterior Envelopes of Buildings V*, pp. 594–601. ASHRAE/DOE/BTECC. Clearwater Beach, Florida.

Melgaard, H., Sadegh, P., Madsen, H., & Holst, J. (1993). Experiment design for grey-box models. In *Preprints of the 12th IFAC World Congress*, Vol. 2, pp. 145–148.

Merton, R. (1971). Optimum consumption and portfolio rules in a continuous-time model. *J. Econom. Theory*, **3**, 373–413. Erratum:ibid. **6** (1973), 213–214.

Milshtein, G. N. (1974). Approximate integration of stochastic differential equations. *Theor. Prob. Appl.*, **19**, 557–562.

Moler, C. & van Loan, C. (1978). Nineteen dubious ways to compute the exponential of a matrix. *SIAM Review*, **20**(4).

Pázman, A. (1986). *Foundations of Optimum Experimental Design*. Mathematics and Its Applications (East European Series). D. Reidel Publ., Dordrecht.

Priestley, M. B. (1981). *Spectral Analysis and Time Series*. Academic Press, London.

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

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*. Wiley, New York.

Ripley, B. D. (1987). *Stochastic Simulation*. Wiley, New York.

Rümelin, W. (1982). Numerical treatment of stochastic differential equations. *SIAM J. Numer. Anal.*, **19**(3), 604 – 613.

Sadegh, P., Melgaard, H., Madsen, H., & Holst, J. (1994). On the usefulness of grey-box information when doing experiment design for system modelling and identification. In *IFAC - SYSID '94*, Vol. 2–145.

Sadegh, P. (1993). Experimental design for grey box models. Master's thesis, IMM, Technical University of Denmark. no. 18/93.

Schwarz, G. (1978). Estimating the dimension of a model. *Ann. Stat.*, **6**, 461–464.

Söderström, T. & Stoica, P. (1989). *System Identification*. Prentice Hall, London.

Sonderegger, R. C. (1977). *Dynamic Models of House Heating Based on Equivalent Thernal Parameters*. Ph.D. thesis, Princeton University, Princeton, N. J. 08540. Report PU/CES 57.

Stoica, P. & Nehorai, A. (1989). On multistep prediction error methods for time series models. *Journal of Forecasting*, **8**, 357–368.

Subba Rao, T. & Gabr, M. M. (1984). *An Introduction to Bispectral Analysis and Bilinear Time Series Models*. Springer-Verlag, Berlin.

Subbarao, K. (1985). Thermal parameters for single and multizone buildings and their determination from performance data. Tech. rep. SERI/TR-253-2617, Solar Energy Research Institute.

Talay, D. (1991). Approximation of upper lyapunov exponents of bilinear stochastic differential systems. *SIAM Journal on Numerical Analysis*, **28**.

Thornton, C. L. & Bierman, G. J. (1980). UDU^{t} covariance factorization for Kalman filtering. In Leondes, C. T. (Ed.), *Control and Dynamic Systems*. Academic Press, New York.

Tong, H. (1990). *Non-linear Time Series, A Dynamical System Approach*. Clarendon Press, Oxford.

Tulleken, H. J. A. F. (1993). Grey-box modelling and identification using physical knowledge and bayesian techniques. *Automatica*, **29**(2), 285–308.

Wahlberg, B. & Ljung, L. (1986). Design variables for bias distribution in transfer function estimation. *IEEE Transactions on Automatic Control*, **AC-31**(2), 134–144.

Walter, E. & Pronzato, L. (1990). Qualitative and quantitative experiment design for phenomenological models - a survey. *Automatica*, **26**(2), 195–213.

Weisberg, S. (1985). *Applied Linear Regression* (2'nd edition). Wiley, New York.

Wouters, P. & Vandaele, L. (1990). *The PASSYS Test Cells*, chap. 1. Commission of the European Communities.

Wu, H., Aquino, C. F., & Chou, G. L. (1983). A 1.6 liter engine and intake manifold dynamic model.. ASME conf. paper No. 83-WA/DSC-39.

Yuan, Z.-D. & Ljung, L. (1985). Unprejudiced optimal open loop input design for identification of transfer functions. *Automatica*, **21**(6), 697–708.

Zarrop, M. B. (1979). *Optimal Experiment Design for Dynamic System Identification*. Springer-Verlag, Berlin.

Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics*. Wiley, New York.

Ph. D. theses from IMM

1. **Rasmus Larsen.** (1994). *Estimation of visual motion in image sequences.* 143 pp.

2. **Jens Moberg Rygaard.** (1994). *Design and optimization of flexible manufacturing systems.* 232 pp.

3. **Niels Christian Krieger Lassen.** (1994). *Automated determination of crystal orientations from electron backscattering patterns.* 136 pp.

4. **Henrik Melgaard.** (1994). *Identification of Physical Models.* 263 pp.