9 11

PARAMETER ESTIMATION APPLIED TO POWER SYSTEM MODELS SUBJECT TO RANDOM DISTURBANCES

by

HANS MULLER

Dipl. El.-Ing. ETH
Swiss Federal Institute of Technology, Zurich
1973

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

THE FACULTY OF GRADUATE SCIENCE

Department of Electrical Engineering

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

June, 1979

Hans Muller

Department of Electrical Engineering

The University of British Columbia
2075 Wesbrook Place
Vancouver, Canada
V6T 1W5

Date 9/7/79

ABSTRACT

For control of an electric generator, it is desirable to approximate the remote part of the power system by a low order model, whose parameters can be estimated from measurements. The proposed model for the remote system consists of a large synchronous machine and a random varying load to account for the dynamic behaviour and for the small fluctuations that are always present.

The maximum likelihood algorithm is a good method for parameter estimation of a noisy system. It can even be used for estimation from measurements alone, without applied disturbance, which could be of great advantage in a power system. The algorithm is derived in a form suitable for efficient digital processing of sampled measurements from a linear continuous system with physical parameters.

The performance of the algorithm is tested by computer simulation of a simplified model for several cases of deterministic and/or stochastic input. It is demonstrated, that good estimates can be found from the output of a disturbed system when no intentional input is applied.

TABLE OF CONTENTS

ACKNOWLEDGEMENT

I would like to thank my thesis supervisor, Dr. M. D. Wvong, as well as Dr. M. Davies for their guidance, advice and encouragement during the research work and writing of this thesis.

My thanks also to Ms. S. Y. Kim for typing the thesis.

My stay in Canada was made possible by an award of the Canada Council. This grant and the financial support of the University of British Columbia are gratefully acknowledged.

NOMENCLATURE

General:

small letter, eg. x or $\underline{x}$	column- vector
x^T	row- vector
capital letter, eg. A	matrix or scalar quantity of power system
A^T	transpose of matrix A
$E[x]$, $\hat{x}$	mean of random vector x
$E[x.x^T]$	covariance matrix of random vector x
$\delta(t)$	Dirac delta
δ_{kl}	Kronecker delta
L, V	likelihood function (scalar)

Time:

t	continuous time
t_k, k	discrete time, $t_k = k.\Delta t$

State space description:

x	state vector
y	(model) output vector
z	(actually measured) output vector
u	control input vector
w	input noise vector
v	measurement noise vector
α	vector of unknown physical parameters
A, B, C, H	system matrices of continuous- time model
F, G, H	system matrices of discrete- time model
Q, R	noise covariance matrices

Kalman filter:

$\hat{x}(-)$, $\hat{x}(+)$	extrapolated and updated state estimate
$P(-)$, $P(+)$	covariance matrix of $\hat{x}(-)$, $\hat{x}(+)$ respectively
e	innovations vector
P_e	innovations covariance matrix
K	Kalman gains

Power system quantities:

all quantities except δ are per unit quantities

E, V	voltage
I	current
P, Q	real and reactive power
X	reactance
X'	transient reactance
T	torque or time constant
T'	transient time constant
M	machine inertia
D	macine damping
δ	angle (in rad) between machine q- axis and synchronous frame of reference
α	angle (in rad) between machine q- axis and bus voltage
$\Delta\omega$	speed deviation from synchronous speed

Indices:

e	electrical
m	mechanical
d	d- axis
q	q- axis
i	local generator i
j	external macine j
t	terminal of local generator
L	load bus
FD	voltage regulator
G	speed governor

I INTRODUCTION

1.1 Dynamic Power System Models

There is continued interest to improve dynamic and transient stability limits in electric power systems. Modern control theory offers new methods to achieve this goal. Linear optimal control designs show great promise [1]. These techniques in turn require a good model for the power system. The interest here is a model for a single generator to be controlled which is connected to a large system.

The classical approach is to model the generator as a machine connected to an infinite bus [2]. This implies that the machine under consideration has no influence upon the dynamic behaviour of the rest of the power system. This assumption is too restrictive, especially for a large generator.

On the other hand, each machine in the entire interconnected system may be modelled individually. The resulting model is then much too large and must be simplified to become tractable.

An intermediate approach is to model the generator under consideration (the "internal" system) in some detail and postulate a simple model for the remaining "external" system. Parameters of this simple model are then identified from measurements taken at the "internal" system generator. This method has two advantages: First, only the dynamics of the external system that interact strongly with the internal system need be represented, and second, the parameters could be estimated by an on-line computer and the model therefore adapted to the actual state of the power network.

The proposed external system usually consists of one (or several) very large synchronous machines [3-7]. Yu, et al. [3] used a small disturbing input signal and estimated the parameters from the system response. It was assumed that there were no disturbances in the external system.

In the following, additional random fluctuations in the external system are considered. The problems treated are the modelling of these disturbances and estimating of parameters in a noisy system.

1.2 Parametric Models for Identification

Basic properties of identification problems are treated by Åström and Eykhoff in [8]. In the following it is assumed that a mathematical model for a power system is derived from a priori knowledge and the identification problem is reduced to the estimation of several unknown parameters. State-space models have been chosen for several reasons: they are easily obtained from differential equations, can describe multi-output systems easily, and form the basis for optimal control.

(a) Deterministic models

If all inputs and outputs of the system can be measured accurately, a general linear state space model is of the form:

$$\underline{x}(k+1) = F(\underline{\alpha}) \cdot \underline{x}(k) + G(\underline{\alpha}) \cdot \underline{u}(k) \tag{1-1}$$

$$\underline{y}(k) = H(\underline{\alpha}) \cdot \underline{x}(k) \tag{1-2}$$

$\underline{\alpha}$ is the vector of unknown parameters and shall not depend on time.

The discrete-time form of the state equation is chosen because the input and output signals shall be sampled and processed by a discrete machine (computer). If a continuous-time model

$$\dot{\underline{x}}(t) = A(\underline{\alpha}).\underline{x}(t) + B(\underline{\alpha}).u(t) \tag{1-3}$$

is derived from differential equations, it can be exactly converted into form (1-1) by integration over one time step, provided that the input $\underline{u}(t)$ changes only at the discrete time points. This is not a severe restriction for a power system.

(b) Stochastic models

The system may have inputs (disturbances) which cannot be measured and the measurements of the outputs may be inaccurate. A state space model for a system with additive noise is:

$$\underline{x}(k+1) = F(\underline{\alpha}) \cdot \underline{x}(k) + G(\underline{\alpha}) \cdot \underline{u}(k) + M(\underline{\alpha}) \cdot \underline{w}(k) \tag{1-4}$$

$$\underline{y}(k) = H(\underline{\alpha}) \cdot \underline{x}(k) + \underline{v}(k) \tag{1-5}$$

$\underline{v}(k)$ and $\underline{w}(k)$ are random variables of input and measurement noise respectively. Additional assumptions like zero mean, independent sequences are usually made.

It should be noted that the model state and output are now random variables too, i.e., $\underline{x}(k)$ and $\underline{y}(k)$ are described by a probability density function.

1.3 Input Signal Requirements

Let $\underline{x}(0)=\underline{0}$. Then identification is only possible if some minimal requirements about the input are met. For many estimators, a sufficient condition of "persistent excitation" was outlined by Åström in [8]. It may be interpreted that the input signal must have some amplitude for every frequency in the band of interest.

Stochastic models have two types of input signals, the control $\underline{u}(k)$ which is deterministic and the noise $\underline{w}(k)$ which is stochastic.

(a) Deterministic inputs

These include disturbances applied for the sake of estimation. Such disturbances are a compromise between what is desired for estimation purposes and what can be realized under physical and operational constraints. For estimation of power system dynamics input signals can be applied at the governor or the voltage regulator. Since any interference into normal operation is undesirable, these signals should be small and of short duration. Yu, et al. [3] proposed a step change of mechanical torque for a short time, the simplest signal with "persistent excitation".

(b) Noise inputs

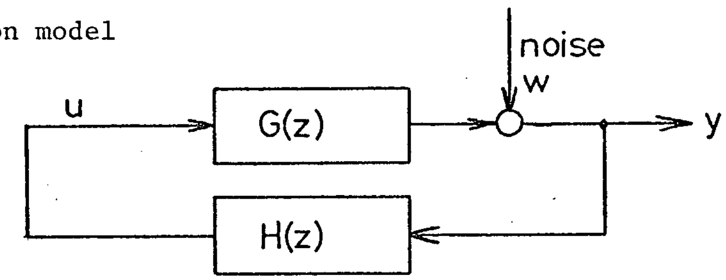
Observed changes in the system under test, which are neither due to applied inputs nor to parameter changes may be attributed to additional stochastic inputs $\underline{w}$. These inputs may be considered a nuisance since they make a correlation of applied input and measured output much more difficult.

But on the other hand, one might ask if the noise alone disturbs the system enough so that identification is possible from output measurements only. A white or colored Gaussian noise sequence for example would meet the requirement of persistent excitation. The less is known about these noise inputs, the system identification gets more difficult, since not only the system parameters $\underline{\alpha}$, but also parameters of the postulated noise process have to be estimated. But this is compensated by the advantage that the estimation process requires no interference of the normal operation of the system at all, an advantage that would count very high for power system identification.

In fact, a few attempts have been made to estimate power system dynamics from normal operating data. Lindahl and Lyung [4] tried to fit several models by processing actual power system measurements and estimated parameters of a very simple 5th order model. They achieved not unreasonable estimates but expressed a desire to introduce additional intentional disturbances. Price et al. [7] tuned simulation programs to produce output signals that were very similar to actual field measurements. They concluded that parameter estimation from on-line measurements is possible, but that a choice of several model structures should be available, based upon knowledge from system planning data.

(c) Effects of noise and feedback

One has to be careful in choosing an input signal that it is not the result of a feedback. An example of what can happen is given in [8] for a transfer function model



An attempt to identify G(z) from measurements would yield the estimate

$$\frac{Y(z)}{u(z)} = \frac{1}{H(z)} \;!$$

For a power system model the excitation voltage and the mechanical torque can be treated as input signals only if the feedback loops of the governor and voltage regulator are open.

1.4 Sources of Input Noise in Power Systems

Field measurements of terminal quantities on a generator (e.g. [7]) show fluctuating signals even if no intentional disturbance is applied. These fluctuations come from changes in the internal or external system and the model of the entire system should take them into account. In this thesis, only fluctuations in the external system power demand will be considered. The simplest way to account for these demand changes is to model the active and reactive power demand as stochastic processes with the steady state power as mean value. Hence all load dynamics should be included in the one-machine model that describes the dynamics of the external system.

Price et al. [5-7] introduced power demand as a stochastic input in their model, whereas Lindahl, et al.[4] introduced a general noise vector $\underline{w}$ with unknown covariance.

The easiest noise process from an estimation standpoint is Gaussian white noise and this rather restrictive assumption is made for most of the following chapters. How a somewhat more realistic noise sequence can be introduced in the power system model is shown in section 2-7.

1.5 Methods for Parameter Estimation

There are a variety of methods available, the problem is to find one which yields an unbiased, reliable, stable estimate with reasonable computing effort.

(a) Parameter estimation as state estimation

For state estimation in linear systems, the Kalman filter is a very efficient and well proven estimator. The parameters $\underline{\alpha}$ in eq. (1-4) can be considered as system states which obey the equation $\underline{\alpha}(k+1) = \underline{\alpha}(k)$. The parameter estimation problem can therefore be viewed as estimation of the states of the augmented model [9]:

$$\begin{bmatrix} \underline{x}(k+1) \\ \underline{\alpha}(k+1) \end{bmatrix} = \begin{bmatrix} F(\underline{\alpha}(k)) & 0 \\ 0 & I \end{bmatrix} \cdot \begin{bmatrix} \underline{x}(k) \\ \underline{\alpha}(k) \end{bmatrix} + \ldots \qquad (1\text{-}6)$$

This method has two drawbacks: The augmented system becomes nonlinear even if the original system was linear and the order becomes higher too. It is not too well suited for estimation of unknown parameters [10], but it should be considered for parameter tracking if very good initial estimates are available.

(b) Least squares methods

Least squares methods involve minimization of a cost which is a quadratic function of some error. The error results from comparison of the model behaviour with the measurements on the real system.

The error is normally so defined that it is linear in the parameters. The cost function can be minimized in one step. These so called "equation error" methods yield efficient algorithms and are easily implemented as recursive estimators. Unfortunately, they can produce strongly biased estimates when noise is present in the system [8].

In "output error" methods, the error is the difference between measured output and model output. This error is not linear in the parameters and nonlinear techniques are applied to optimize the cost function. This method was used by Yu, et al. [3] for a deterministic system. It can be generalized for the case with noise and is related to the maximum likelihood method (see section 2-5).

(c) Maximum likelihood method

The maximum likelihood method is designed for stochastic system models where the outputs are random variables. It basically adjusts the model parameters in a way that maximizes the probability that the model delivers the measured output. It is reported to yield stable, unbiased estimates [8] and has been shown useful for power system estimation [4-7, 11].

The maximum likelihood method works for systems with several outputs and also for systems with only noise inputs. Although it is numerically expensive, it can be implemented on a modern large computer.

2. THE MAXIMUM LIKELIHOOD ALGORITHM FOR PARAMETER ESTIMATION

In this chapter, the maximum likelihood algorithm for parameter estimation is developed. The major component of the estimator is a linear optimal filter (Kalman filter). The filter presented is useful for digital processing of sampled measurements from a continuous process. In the case of a general white input noise, the model and filter should be rearranged with the filter gain as unknown stochastic parameters to insure uniqueness. Alternatively, an input-noise model of low order may be preferable.

A comparison with a least squares algorithm is made. Some methods for nonlinear function optimization are presented. The system model is extended to include correlated noise input.

2.1 State Space Description and Likelihood Function

Let the physical system be described by the following equations:

$$\dot{x}(t) = A(\alpha) \; . \; x(t) + B(\alpha) \; . \; u(t) + C(\alpha) \; . \; w(t) \qquad (2\text{-}1)$$

$$y(t_k) = H(\alpha) \; . \; x(t_k) + v(t_k) \qquad (2\text{-}2)$$

with the vectors:

x: state u: control input y: model output

w: plant or input noise v: measurement noise

α: unknown physical parameters

The mixed continuous (2-1) and discrete (2-2) form indicates that the model is derived from differential equations but measurements are sampled and processed in discrete time. w is a Gaussian white noise process with zero mean and constant covariance Q, i.e.:

$$E[w(t)] = 0 \quad E[w(t) . w^T(\tau)] = Q . \delta(t-\tau) \qquad (2\text{-}3)$$

v is a Gaussian white noise vector with zero mean and constant covariance R, i.e.:

$$E[v(k)] = 0 \quad E[v(k) . v^T(\ell)] = R . \delta_{k\ell} \qquad (2\text{-}4)$$

where the index k denotes time t_k.
In addition to α, Q and R may be unknown too.

With these assumptions for v and w, the probability density function of the output pr(y(k)) can be calculated for all k as a function of α, Q and R from Eqs. (2-1, 2-2).

The likelihood function is the value of the joint probability density for the actually measured output sequence $pd_{y(N),y(N-1),\ldots,y(1)}(z(N),z(N-1),\ldots \ldots,z(1)|\alpha,Q,R)$ where $z(1)\ldots z(N)$ denote the available output measurements. The maximum likelihood method attempts to find those parameter estimates $\hat{\alpha}$, $\hat{Q}$, $\hat{R}$ for which the likelihood function is maximized.

Calculation of the log likelihood function [10]

The joint probability density function is computed recursively by application of Bayes' rule. With the abbreviation

$$Y(k) \hat{=} [y(k),y(k-1),\ldots,y(1)] \text{ and } Z(k) \hat{=} [z(k),z(k-1),\ldots,z(1)] \quad (2\text{-}5)$$

$$pd_{Y(k)}(Z(k)|\alpha,Q,R) = pd_{y(k)}(z(k)|Z(k-1);\alpha,Q,R) \;.\; pd_{Y(k-1)}(Z(k-1)|\alpha,Q,R) \quad (2\text{-}6)$$

If L denotes the negative logarithm of the likelihood function, Eq. (2-6) becomes

$$-L(k;\alpha,Q,R) = -L(k-1;\alpha,Q,R) + \ln \, pd_{y(k)}(z(k)|Z(k-1);\alpha,Q,R) \tag{2-7}$$

For L(0)=0 and with the recursion carried out

$$L(N;\alpha,Q,R) = -\sum_{k=1}^{N} \ln \, pd_{y(k)}(z(k)|Z(k-1);\alpha,Q,R) \tag{2-8}$$

Minimizing L will maximize the likelihood, since the logarithm is a monotonic function. It remains to evaluate the conditional probability density of y(k). Because of the Gaussian assumption for the noises w and v, this density is also a normal distribution.

The normal distribution for a random variable is

$$pd_y(z) = (2.\pi.\sigma^2)^{-\frac{1}{2}} \, . \, \exp(-e^2/2\sigma^2) \tag{2-9}$$

where $e=z-\hat{y}$, $\hat{y}=E[y]$, $\sigma^2= E[(y-\hat{y})^2]$

The corresponding multivariable normal distribution for a random vector is

$$pd_{y(k)}(z(k)|Z(k-1)) = [(2.\pi)^m . \det P_y(k|k-1)]^{-\frac{1}{2}} . \exp[-\frac{1}{2} e^T(k) \, . \, P_y^{-1}(k|k-1).e(k)] \tag{2-10}$$

where m = dimension of the vector y

$$e(k) = z(k) - \hat{y}(k|k-1)$$

$$\hat{y}(k|k-1) = E[y(k)|Z(k-1);\alpha,Q,R]$$

$$P_y(k|k-1) = E[(y(k)-\hat{y}(k))(y(k)-\hat{y}(k))^T |Z(k-1);\alpha,Q,R]$$

The conditional mean and the covariance of y(k) are provided by an <u>optimal linear filter</u> (Kalman filter) that processes the measurements z(1) ... z(k-1) for the specified values of α, Q, R.

2.2 The Continuous-Discrete Kalman Filter

The Kalman filter for the model (2-1, 2-2) is a combination of the discrete and continuous filters described by Gelb [9].

(a) Propagation in the absence of measurements

Between t_{k-1} and t_k no measurements are available and the state estimate $\hat{x}(t)$ and its error covariance P(t) are simply extrapolated:

$$\dot{\hat{x}}(t) = A \,.\, \hat{x}(t) + B \,.\, u(t) \quad (t_{k-1}<t<t_k) \tag{2-11}$$

$$\dot{P}(t) = A \,.\, P(t) + P(t).A^T + C.Q.C^T (t_{k-1}<t<t_k) \tag{2-12}$$

(b) Update with new measurement

At time t_k a new measurement, $z(t_k)$, is available and is used to update state estimate and covariances. There is a discontinuity at this point and the indices t_k^- and t_k^+ (or simply k- and k+) are used to indicate the quantities before and after the updating takes place.

$\hat{y}(k-) = H \cdot \hat{x}(k-)$	output estimate	(2-13)
$e(k) = z(k) - \hat{y}(k-)$	innovations	(2-14)
$P_e(k) = H \cdot P(k-) \cdot H^T + R$	innovations covariance	(2-15)
$\hat{x}(k+) = \hat{x}(k-) + K(k) \cdot e(k)$	updated state estimate	(2-16)
$P(k+) = [I - K(k) \cdot H] \cdot P(k-)$	updated error covariance	(2-17)
$K(k) = P(k-) \cdot H^T \cdot [H \cdot P(k-) \cdot H^T + R]^{-1}$	Kalman gains	(2-18)

$\hat{y}(k-)$ and $P_e(k)$ are the required mean and covariance $\hat{y}(k|k-1)$ and $P_y(k|k-1)$ in the likelihood formula (2-10).

Substitution of (2-10) into (2-8) yields

$$L(N;\alpha,Q,R) = \frac{1}{2}\sum_{k=1}^{N} m\cdot\pi + \ln \det P_e(t) + e^T(k)\cdot P_e^{-1}(k)\cdot e(k)] \qquad (2\text{-}19)$$

The calculation of the likelihood value can be done by a slightly modified Kalman filter.

Solution of the Differential equations (2-11, 2-12)

It is possible to integrate both $\hat{x}(t)$ and $P(t)$ numerically, for example by the Runge-Kutta method, but the computational burden would be excessive. Since the model parameters are all time-invariant, Eq. (2-11) and (2-12) may be converted into difference equations.

$$\dot{\hat{x}}(t) = A\cdot\hat{x}(t) + B\cdot u(t) \quad t_{k+1}<t<t_k \ , \ \hat{x}(t_{k-1}^+) \text{ given} \qquad (2\text{-}11)$$

the superposition integral is

$$\hat{x}(t_k^-) = \Phi(t_k-t_{k-1})\cdot\hat{x}(t_{k-1}^+) + \int_{t_{k-1}^+}^{t_k^-} \Phi\,(t_k-\tau)\cdot B\cdot u(\tau)d\tau$$ with the transition matrix $\Phi(\tau) = \exp(A\cdot\tau)$. If the input u is allowed to change its value only at the time points t_k , k=1 ... N (see section 1-2), then the constant term $B\cdot u(\tau)=B\cdot u(t_{k-1}^+)$ can be taken out of the integral.

With $\quad \Delta t \hat{=} t_k-t_{k-1} \qquad (2\text{-}20)$

$$F(\alpha) \hat{=} \Phi(\Delta t,\alpha) = \exp(A(\alpha)\cdot \Delta t) \qquad (2\text{-}21)$$

$$G(\alpha) \hat{=} \int_0^{\Delta t} \exp[A\cdot(\Delta t-\tau)]\cdot d\tau\cdot B = A^{-1}(\alpha)\cdot[F(\alpha)-I]\cdot B(\alpha) \qquad (2\text{-}22)$$

the integrated version of (2-11) is

$$\hat{x}(t_k^-) = F(\alpha)\cdot\hat{x}(t_{k-1}^+) + G(\alpha)\cdot u(t_{k-1}^+) \qquad (2\text{-}23)$$

For the coveriance equation (2-12) the same approach as for the Matrix Ricatti equation [9, p.136] is taken. For the equation $\dot{P} = A \cdot P + P \cdot A^T + CQC^T$ (2-12) the transformations $\lambda \hat{=} P \cdot \mu$ and $\dot{\mu} \hat{=} -A^T \cdot \mu$ define a linear system

$$\begin{bmatrix} \dot{\mu} \\ \dot{\lambda} \end{bmatrix} = \begin{bmatrix} -A^T & 0 \\ C \cdot Q \cdot C^T & A \end{bmatrix} \cdot \begin{bmatrix} \mu \\ \lambda \end{bmatrix}$$

with the solution

$$\begin{bmatrix} \mu(\Delta t) \\ \lambda(\Delta t) \end{bmatrix} = \begin{bmatrix} (F^{-1})^T & 0 \\ T & F \end{bmatrix} \cdot \begin{bmatrix} \mu(0) \\ \lambda(0) \end{bmatrix}$$

where

$$\begin{bmatrix} (F^{-1})^T & 0 \\ T & F \end{bmatrix} = \exp\left(\Delta t \cdot \begin{bmatrix} -A^T & 0 \\ C.Q.C^T & A \end{bmatrix} \right) \tag{2-24}$$

and for P

$$P(t_k^-) = T \cdot F^T + F \cdot P(t_{k-1}^+) \cdot F^T \tag{2-25}$$

There are many methods to evaluate the matrix exponentials numerically, see section 4-3.

2.3 Steady State of the Kalman Filter

The calculation of one likelihood value $L(N;\alpha, Q,R)$ as developed in the previous sections involves two steps

(1) For given values of α, Q, R and $P(0) = E[x(0) \cdot x^T(0)]$, the system matrices $F(\alpha)$, $G(\alpha)$, $H(\alpha)$, the Kalman gains $K(t_k)$, k=1...N and the innovations covariance $P_e(t_k)$, k=1...N are calculated.

(2) The measurements $u(0) \ldots u(N-1)$; $z(1) \ldots z(N)$ are processed by the Kalman filter to yield the innovation sequence $e(1) \ldots e(N)$ and L. Since both the system dynamics F, G, H and the noise statistics Q, R are assumed independent of time, the Kalman filter will reach a statistical steady state, i.e., the matrices $P(t_k^+)$, $P(t_k^-)$, $P_e(t_k^-)$, $K(t_k)$ converge to constant matrices $P(+)$, $P(-)$, P_e, K. The estimation algorithm requires relatively long data sequences (> 100 points), the time dependent covariance and gain matrices can therefore be replaced by the steady-state matrices, with the additional advantage, that dependence upon the unknown initial covariance $P(0)$ is removed.

The steady state value could be obtained by solving the matrix equation $P(t_k^-) = P(t_{k-1}^-)$. However, this system of nonlinear algebraic equations is difficult to solve. It is numerically more convenient to calculate the Kalman gains as shown in the first step above with an arbitrary matrix $P(0)$ and use the values to which the gains converge.

For the power system model considered in section 3-5 and the choice $P(0)=0$, convergence to 5 digits was usually obtained in 5 to 10 steps, an indication that statistical steady-state can safely be assumed.

Summary of Equations for the Likelihood Function

System model (process):

$$\dot{x}(t) = A(\alpha) \cdot x(t) + B(\alpha) \cdot u(t) + C(\alpha) \cdot w(t) \tag{2-1}$$

$$z(k) = H(\alpha) \cdot x(t_k) + v(k) \tag{2-2}$$

$$E[w(t) \cdot w^T(\tau)] = Q \cdot \delta(t-\tau) \tag{2-3}$$

$$E[v(k) \cdot v^T(\ell)] = R \cdot \delta_{k\ell} \tag{2-4}$$

Kalman Filter:

$$\hat{x}(k-) = F\cdot\hat{x}(k-1+) + G\cdot u(k-1+) \qquad (2\text{-}23)$$

$$e(k) = z(k) - H\cdot\hat{x}(k-) \qquad (2\text{-}13/14)$$

$$\hat{x}(k+) = x(k-) + K\cdot e(k) \qquad (2\text{-}16)$$

likelihood function (constant term mπ neglected):

$$L(N;\ \alpha,\ Q,R) = \frac{1}{2}\{N\cdot \ell n \det P_e + \sum_{k=1}^{N} e^T(k)\cdot P_e^{-1}\cdot e(k)\} \qquad (2\text{-}19)$$

Steady state matrices:

$$K = \lim_{k\to\infty} K(k)\ , \qquad P_e = \lim_{k\to\infty} P_e(k) \text{ from}$$

$$P(k-) = F\cdot P(k-1+)\cdot F^T + T\cdot F^T\ ; \quad P(0+) = 0 \qquad (2\text{-}25)$$

$$P_e(k) = H\cdot P(k-)\cdot H^T + R \qquad (2\text{-}15)$$

$$K(k) = P(k-)\cdot H^T\cdot P_e^{-1}(k) \qquad (2\text{-}18)$$

$$P(k+) = [I-K(k)\cdot H]\cdot P(k-) \qquad (2\text{-}17)$$

Remark: This form of the Kalman filter is valid also in the case of perfect measurements (R=0), although numerical difficulties may arise when inverses are calculated.

Continuous-discrete relationship:

$$F = \exp(\Delta t\cdot A(\alpha)) \qquad G = A^{-1}\cdot(F - I)\cdot B \qquad (2\text{-}21,\ 2\text{-}22)$$

$$\begin{bmatrix} (F^{-1})^T & 0 \\ T & F \end{bmatrix} = \exp\left(\Delta t\cdot\begin{bmatrix} -A^T & 0 \\ C.Q.C^T & A \end{bmatrix}\right) \qquad (2\text{-}24)$$

By use of the identify

$$x^T\cdot A\cdot y = \text{trace}\ (A\cdot y\cdot x^T)$$

the time independent matrix P_e can be taken out of the summation in the likelihood formula. Hence

$$L(N;\alpha,Q,R) = N\cdot \ell n \det P_e + \mathrm{trace}(P_e^{-1} \sum_{k=1}^{N} e(k)\cdot e^T(k)) \quad (2\text{-}27)$$

2.4 Maximum Likelihood Estimates

The existence of optimal parameter estimates is now investigated.

Uniqueness of optimal parameters

The major drawback of state space models is their non-uniqueness; many combinations of the matrices A,B,C,H account for the same input-output sequence. For multivariable systems even canonical forms are not unique [11]. For a physically derived model with a small set of parameters α, this need not be a big problem. Care should be exercised in the choice of parameters, so that the transfer functions $H\cdot (sI-A)^{-1}B$ and $H\cdot(sI-A)^{-1}C$ are unique.

Both Åström [11] and Kashyap [12] point out, that in general the noise statistics (Q and R) cannot be determined by the likelihood method. Different noise covariances may give the same values for K and P_e and therefore the same likelihood L. Åström suggests that L be minimized as a function of α, K and P_e.

The Innovations Representation [8,11]

The equations of Kalman filter and system model can be combined to the following model:

$$\hat{x}(t_k^-) = F(\alpha)\cdot\hat{x}(t_{k-1}^-) + G(\alpha)\cdot u(t_{k-1}) + K'\cdot e(t_{k-1}) \quad (2\text{-}28)$$

$$z(t_k) = H(\alpha)\cdot\hat{x}(t_k) + e(t_k) \quad (2\text{-}29)$$

where $K' = F.K$ and $e(t_k)$, $k=1..N$ is a sequence of independent vectors with zero mean and covariance P_e. These equations can be solved for the innovations e if α, K', u and z are given. The likelihood function is

$$L(N;\alpha,K';P_e) = \frac{1}{2}\{N\cdot\ln\det P_e + \mathrm{trace}(P_e^{-1}\cdot\sum_{k=1}^{N} e(k)\cdot e^T(k))\} \qquad (2\text{-}30)$$

The minimization with respect to P_e is done analytically. Setting the derivatives of L to each element of P_e to zero gives the condition

$$\hat{P}_{e\ opt} = \frac{1}{N}\sum_{k=1}^{N} e(k)\cdot e^T(k) \qquad (2\text{-}31)$$

The maximum likelihood algorithm is then restated as:

$$\text{minimize } V(N;\alpha,K') = \det\sum_{k=1}^{N} e(k)\cdot e^T(k) \qquad (2\text{-}32)$$

$$e(k) = z(k) - H(\alpha)\cdot\hat{x}(k) \qquad (2\text{-}33)$$

$$\hat{x}(k) = F(\alpha)\cdot\hat{x}(k-1) + G(\alpha)\cdot u(k-1) + K'e(k-1) \qquad (2\text{-}34)$$

$$\hat{x}(0) = 0, \qquad e(0) = 0$$

This is very similar to a minimum variance estimator; the relationship is shown in the next section.

Kashyap [12] provides a proof of convergence and the necessary conditions for estimation from an innovations representation. However, he uses a different type of input-output description to avoid the ambiguities of the state-space model. The conclusions of [12] can be interpreted in the following way:

Under the assumptions of

- zero mean, Gaussian input noise with a certain covariance
- zero mean, independent Gaussian measurement noise

- stable feedback system and linear optimal filter (i.e. eigenvalues of both F and F-K'H inside unit circle)

the maximum likelihood estimates of both the parameters in F and H and the filter gains K will converge to the true values as the number of samples N goes to infinity (NB: This includes the case with no deterministic inputs).

Choise of reference model

For the theoretical reasons mentioned, the innovations description should be used for estimation of a general model. This was done by Lindahl [4], who made no special assumptions about the manner in which the input noise enters the system.

On the other hand, if there is a simple physical description of the source of the input noise, then the noise covariance Q may contain only a few unknown entries. The simple power system model of section 3-5 contains only one stochastic parameter (the load demand variance), but at least six Kalman gains. In this case, the original function $L(\alpha,Q,R)$ should be optimized, since it involves fewer parameters. Price et al. [5] also choose this approach for their model.

2.5 Least Squares Algorithms

(a) Single output systems

For a physically derived model with measurement noise the output error method is used. The cost function is

$$J(\alpha) = \sum_{k=1}^{N} e^2(k/\alpha) , \qquad e(k/\alpha) = z(k)-\hat{y}(k/\alpha) \qquad (2\text{-}35)$$

The output error or residual e is the difference between the measured output z and the output $\hat{y}$ of a reference model. Nonlinear optimization methods are used to minimize $J(\alpha)$.

Convergence properties of the estimates depend on the reference model that provides the $\hat{y}(k)$. If the reference model is of the form

$$\hat{x}(k) = F(\alpha)\cdot\hat{x}(k-1) + G(\alpha)\cdot u(k-1) \tag{2-36}$$

$$\hat{y}(k) = H(\alpha)\cdot\hat{x}(k) \tag{2-37}$$

which does not include any noise, but the measurements are created by a process disturbed by noise at the input (w(k)), then the residuals e(k) will be a correlated sequence. In this case the least squares estimate is biased [8].

For transfer function models with noise, the least squares method is generalized by pre-filtering the measurements u(k) and z(k) in such a way that the residuals e(k) become uncorrelated. For a state space model, this is accomplished by the Kalman filter. If the innovations model (2-33, 2-34) is used as reference instead of (2-36,2-37), then the residuals are the uncorrelated innovations.

In fact, for a single output system the likelihood function (2-30) is

$$V(\alpha,k) = \det \sum_{k=1}^{N} e(k)e^T(k) = \sum_{k=1}^{N} e^2(k) = J(\alpha,K) \tag{2-38}$$

In the case of a single output, the <u>maximum likelihood estimate</u> and the <u>least squares estimate</u> (for the innovation model) are accordingly indentical. In addition, it is also the <u>minimum variance estimate</u>, since $J(\alpha,K)$ is an unbiased approximation of the variance of e.

(b) Multiple outputs

The least squares cost function for several outputs is

$$J(\alpha) = \sum_{k=1}^{N} e^T(k/\alpha)\cdot W\cdot e(k/\alpha) \quad ; \quad e(k/\alpha)=z(k)-\hat{y}(k/\alpha) \qquad (2\text{-}39)$$

with an arbitrary weighting matrix W.

If the e(k) are independent Gaussian vectors (from an innovations model) then it can be shown that the optimal choice for W is the inverse of the covariance matrix and for this choice the least squares estimate is equal to the minimum variance estimate with the cost function

$$J'(\alpha) = \text{trace} \sum_{k=1}^{N} e(k)\cdot e^T(k) = \sum_{k=1}^{N} \sum_{i} e_i^2(k) \qquad (2\text{-}40)$$

this is different from the likelihood function which involves the determinant of the covariance of e. Kashyap [12] points out that the likelihood function is the correct criterion and that there are cases where estimation with Eq. (2-40) does not yield the true parameters.

2.6 Optimization Methods

With a minimum for the likelihood function guaranteed and the necessary equations to compute function values available, the problem of how to find the minimum value still remains. A great number of nonlinear optimization methods exist, but it is not easy to find one that is both reliable and efficient. All the methods discussed here are described in Adby and Dempster [13]; some of them are available at UBC as subroutines, see Manual UBC NLP [14].

(a) Search methods

These methods attempt to find an optimum by using function evaluations only. Problems arising from computation of derivatives are avoided, but a large number of function evaluations are needed. A simple method is the nonlinear simplex technique, a more sophisticated one is Powell's method of conjugate directions.

(b) Gradient methods

Gradient methods are based on a first order Taylor expansion for the objective function

$$V(\alpha_o+\Delta\alpha) = V(\alpha_o)+g^T(\alpha_o)\Delta\alpha, \quad g = \frac{\partial L}{\partial \alpha} \qquad (2\text{-}41)$$

They are normally superior to search methods for functions with continuous derivatives. Examples are steepest descent and conjugate gradient methods.

A second order Taylor expansion involving the Hessian matrix

$$V(\alpha^o+\Delta\alpha) = V(\alpha^o)+g^T.\Delta\alpha+ \frac{1}{2}\Delta\alpha^T. \ H \cdot\Delta\alpha \qquad (2\text{-}42)$$

leads to Newton's method with the update formula

$$\Delta\alpha = -H^{-1}. \ g \qquad (2\text{-}43)$$

The direct use Newton's method is limited, because the Hessian is usually very expensive to evaluate and the method is prone to divergence for initial values not near the optimum. But there are many approximations, known as quasi-Newton methods, that seek to approximate theHessian by a positive-definite matrix using only derivatives of first order. For least squares problems the Gauss-Newton method is often applied, with several possible modifications. Another group consists of the variable metric methods, the best known of which is the DFP (Davidon-Fletcher-Powell) algorithm.

Bard [15] compared several of the gradient methods for several static parameter estimation problems and found that the Gauss-Newton methods usually performed best.

(c) Evaluation of gradients

For the innovations model (2-31, 2-32) and the likelihood function

$$V'(\alpha,K) = \frac{1}{2}\,\ell n\, V(\alpha,k) = \frac{1}{2}\,\ell n \det S; \; S = \sum_k e(k)\cdot e^T(k) \qquad (2\text{-}44)$$

the derivative to the physical parameter α_i is given by

$$\partial V'/\partial\alpha_i = \frac{1}{2}\,\mathrm{tr}\left(\frac{\partial s}{\partial\alpha_i}\cdot S^{-1}\right) = \sum_k \frac{\partial e^T(k)}{\partial\alpha_i}\cdot S^{-1}\cdot e(k) \qquad (2\text{-}45)$$

and the sensitivity equations

$$\partial e/\partial\alpha_i = -\partial H/\partial\alpha_i \cdot \hat{x}(k) + H \cdot \partial\hat{x}(k)/\partial\alpha_i \qquad (2\text{-}46)$$

$$\partial\hat{x}/\partial\alpha_i = \partial F/\partial\alpha_i\cdot\hat{x}(k-1)+F\cdot\partial\hat{x}(k-1)/\partial\alpha_i+\partial G/\partial\alpha_i\cdot u(k-1) \qquad (2\text{-}47)$$

The sensitivity matrices $\partial F/\partial\alpha_i$, $\partial G/\partial\alpha_i$ are finally obtained from $\partial A/\partial\alpha_i$, $\partial B/\partial\alpha_i$.

For a physical power system model as developed in Chapter 3, the analytical derivatives of the system matrices w.r.t. the parameters become very complicated.

A computationally simpler approach is to perturb each parameter in turn by a small amount $\Delta\alpha_i$ and evaluate the gradient from the approximations

$$\partial V/\partial\alpha_i \approx [V(\alpha+\Delta\alpha_i) - V(\alpha)]/\Delta\alpha_i \qquad (2\text{-}48)$$

Each perturbation $\Delta\alpha_i$ must be small enough that the linear approximation is true, but large enough that roundoff errors are kept small.

A gradient evaluation requires as many function evaluations as there are unknown parameters.

Choice of an optimization method

The theoretical advantage of the gradient methods seems questionable in view of the computational effort required to compute the gradient. The main reason for this is that the model derived by physical principles tends to have only a small number of parameters, which enter the model in a complex way.

The option to experiment with several methods on the actual model should be left open. At UBC this is easily possible by working with the Nonlinear Optimization Monitor [14], which offers a choice of optimization methods.

From the limited experience that was gained by working with the model developed in section 3-5, if seems that Powell's method of conjugate directions performs best for this type of model.

2.7 Correlated Input Noise

The model proposed in section 2-1 may be inadequate for two reasons:

- The input noise w is not white but band limited and would better be modelled as a first order Gauss-Markov process:

$$\dot{w}(t) = A_w \cdot w(t) + s(t) \quad , \quad s(t) \text{ white noise} \tag{2-49}$$

- The input noise w may influence the output directly, so that Eq. (2-2) is replaced by

$$y(t) = H(\alpha)\cdot x(t)+D(\alpha)\cdot w(t)+v(t) \tag{2-50}$$

The resulting complications for the estimator can be resolved by including w in an augmented state vector [6]:

$$\begin{bmatrix} \dot{x}(t) \\ \\ \dot{w}(t) \end{bmatrix} = \begin{bmatrix} A(\alpha) & C(\alpha) \\ \\ 0 & A_w \end{bmatrix} \cdot \begin{bmatrix} x(t) \\ \\ w(t) \end{bmatrix} + \begin{bmatrix} B(\alpha) \\ \\ 0 \end{bmatrix} \cdot u(t) + \begin{bmatrix} 0 \\ \\ s(t) \end{bmatrix} \tag{2-51}$$

$$y(t_k) = [H(\alpha) \;\; D(\alpha)] \cdot \begin{bmatrix} x(t_k) \\ \\ w(t_k) \end{bmatrix} + v(t_k) \tag{2-52}$$

This augmented stated space model has the same stochastic properties as the original one. The elements of the matrix A_w may be considered additional unknown parameters or may be assigned empirical values.

3. DYNAMIC POWER SYSTEM MODEL

A two machine equivalent with random load demand is used as a dynamic power system model. A general method to obtain a linearized state space model from a system of mixed algebraic and differential equations is presented and applied for the two-machine case. A simplified version of the model is used for testing the estimation algorithm.

3.1 Structure of the Model

The following model is used for parameter estimation:

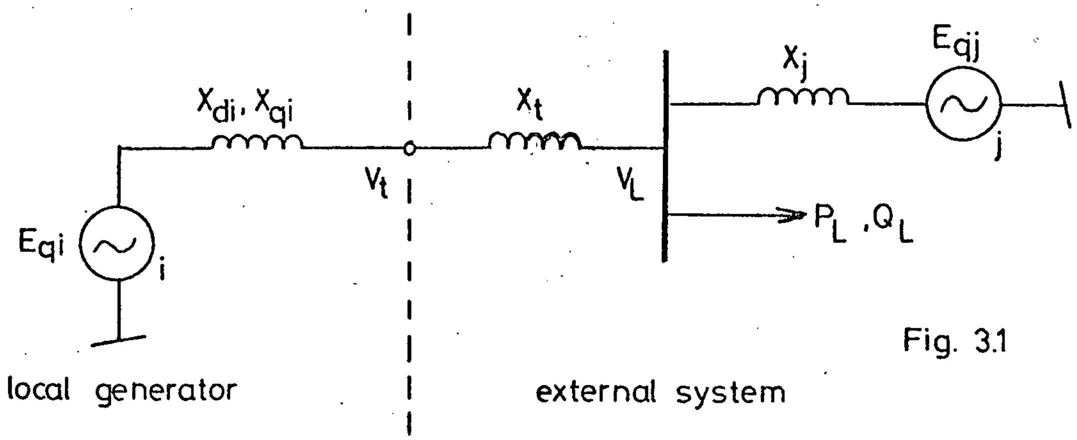


Fig. 3.1

- local generator (i):

This is the generator for which a dynamic bus model shall be identified. The generator is described as a salient pole machine of 3rd order (state variables δ_i, ω_i, E'_{qi}) with an additional voltage regulator and governor, each of first order (state variables E_{fdi}, T_{mi}). All parameters are known. Measurements can be made at this local generator only. They include the terminal quantities V_t, P_t, Q_t and the rotor speed ω_i.

- external system:

The traditional concept of an infinite bus is replaced by a simple external system consisting of:

- a large external round-rotor synchronous machine (j) which represents the dynamics of the external system. A third order description includes the equations of motion and armature reaction (state variables δ_j, ω_j, E'_{qj})

  unknown parameters:

  M_j - total external inertia

  D_j - total external damping

  X_j - steady-state reactance of machine j

  X'_j - transient reactance

  T'_{0j} - transient open circuit time constant

- a bus load $P_L + jQ_L$. No dynamics are included for this load, but the load demand is assumed to vary randomly in the following way: $P_L = P_{LO} + \Delta P_L$, where P_{LO} is the steady-state load and ΔP_L is a Gaussian white-noise process.

- the network interconnection, which consists of the unknown transmission line reactance X_t between the machine terminal and the fictitious bus. Only algebraic equations are considered. These are obtained from phasor diagrams for synchronous speed.

Depending on the size of the load P_{LO} , Q_{LO} , the external machine is operating as a motor or a generator. The sign convention for a generator is always employed.

3.2 Derivation of the State Equations

The following approach was chosen because it allows both simplification of the resulting model as well as generalization for more machines.

- The desired model is linear, valid for small transient deviations from steady state. It is obtained by a first order approximation of the nonlinear equations about the steady state. The steady state is evaluated separately, it depends upon the unknown parameters (see section 3-6).

From the physical model, two types of equations are obtained: (1) state equations of one machine with respect to the load bus (section 3-3): They can be written in the form $\dot{x}=f(x,u,r)$ where x are the states $(\delta, \omega, E_q', \ldots)$, u are the control inputs and r are intermediate variables of the load bus (δ_L, V_L). The linearized one machine model developed in section 3-3 is very similar to the one of De Mello and Concordia [2] with two differences:

- the voltage V_L of the load bus is variable
- the machine angle δ_i is different from the angle α_i between machine q-axis and bus voltage $(\alpha_i=\delta_i-\delta_L)$

The output equations are of the form $y=h(x,r,v)$ where v stands for the measurement noise.

The linearized equations of all machines are combined and one machine angle can be eliminated.

(2) Algebraic equations for the connecting network (load bus, section 3-4): Energy balance conditions at the load bus are expressed by equations $g(x,r,w)=0$ where w stands for the random loads P_L, Q_L.

In order to obtain a single state space model, the equations g=0 are solved for the intermediate variables r and substituted into the state equations. In the nonlinear form, this constitutes a load-flow problem which can be solved by iteration only. However, in the linearized case $\Delta g(\Delta x, \Delta r, \Delta w)=0$ can be solved directly by Gaussian elimination. For the very simple model of section 3-5 the substitution is done analytically.

3.3 Dynamic Equations for One Machine

- equations of motion

$$M \cdot \Delta\dot{\omega} + D \cdot \Delta\omega = T_m - T_e \tag{3-1}$$

$$\dot{\delta} = \omega_o \cdot \Delta\omega \tag{3-2}$$

all quantities in per unit, except δ[rad] and ω_o=120π rad/sec

$M = \frac{J \cdot \omega_o^2}{P_{base}}$ inertia constant

D damping constant

T_m mechanical torque

$T_e \approx P_e = V_d \cdot I_d + V_q \cdot I_q$ electrical torque

$\Delta\omega = \frac{\omega_{rot} - \omega_o}{\omega_o}$ speed deviation

δ angle between rotor axis and synchronous frame of reference

- electrical equations

assumptions:

- 3 windings: d,q,f
- resistance in stator winding negligible
- flux changes in d and q-axis negligible
- speed deviation negligible

from Kimbark [16], p. 73 :

$$\psi_d = M_f \cdot I_f - L_d \cdot I_d$$

$$\psi_q = -L_q \cdot I_q$$

$$\psi_f = L_{ff} \cdot I_f - \frac{3}{2} M_f \cdot I_d$$

$$V_d = -\omega_o \cdot \psi_q$$

$$V_q = \omega_o \cdot \psi_d$$

$$E_f = R_f \cdot I_f + \dot{\psi}_f$$

with $X_d \hat{=} \omega_o \cdot L_d$, $X_q \hat{=} \omega_o \cdot L_q$, $T'_{do} \hat{=} \frac{L_{ff}}{R_f}$ $X'_d = X_d - \frac{3}{2} \frac{\omega_o M_f^2}{L_{ff}}$

$$E_q \hat{=} \omega_o \cdot M_f \cdot I_f \quad E'_q \hat{=} \frac{\omega_o . M_f}{L_{ff}} \cdot \psi_f \quad E_{FD} \hat{=} \frac{\omega_o M_f}{R_f} \cdot E_f$$

the following equations are obtained:

$$T'_{do} \cdot \dot{E}'_q + E'_q = E_{FD} - (X_d - X'_d) \cdot I_d \tag{3-3}$$

$$V_d = X_q \cdot I_q \tag{3-4}$$

$$V_q = E_q - X_d \cdot I_d = E'_q - X'_d \cdot I_d \tag{3-5}$$

In the following a linearized model with the load bus as reference is developed.

Power transfer between two buses:

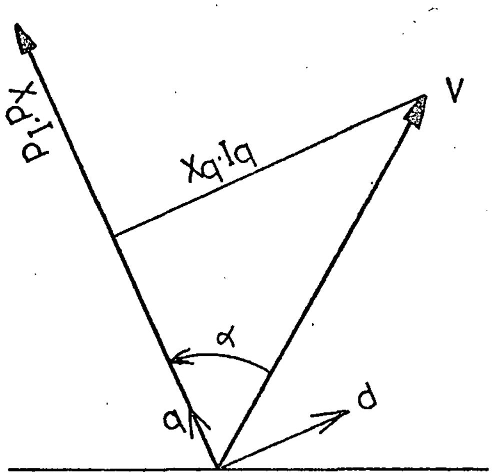


Fig. 3.2

This simplified phasor diagram of Fig. 3.2 is used in the following. Equations for both machines (i) and (j) are obtained through the substitutions:

Machine (i): $E=E'_{qi}$, $V=V_L$, $X_d=X'_{di}+X_t$, $X_q=X_{qi}+X_t$,

$\alpha=\alpha_i=\delta_i-\delta_L$, $I_d=I_{di}$, $I_q=I_{qi}$

Machine (j): $E=E'_{qj}$, $V=V_L$, $X_d=X'_j$, $X_q=X_j$

$\alpha=\alpha_j=\delta_j-\delta_L$, $I_d=I_{dj}$, $I_q=I_{qj}$

incoming power at the load bus V:

$$S = P+jQ = \bar{V}\cdot \bar{I}^* = (V_d+jV_q)\cdot(I_d-jI_q) \tag{3-6}$$

$$P = V_d\cdot I_d + V_q\cdot I_q \tag{3-7}$$

$$Q = -V_d\cdot I_q + V_q\cdot I_d \tag{3-8}$$

$$V_d = V\cdot \sin\alpha = X_q\cdot I_q$$

$$V_q = V\cdot \cos\alpha = E-X_d\cdot I_d$$

Substitution of I_d and I_q yields:

$$P = \frac{V \cdot E \cdot \sin\alpha}{X_d} + \frac{1}{2} V^2 \cdot \sin 2\alpha \left(\frac{1}{X_q} - \frac{1}{X_d} \right) \tag{3-9}$$

$$Q = \frac{V \cdot E \cdot \cos\alpha}{X_d} - V^2 \cdot [\sin^2\alpha \cdot \left(\frac{1}{X_q} - \frac{1}{X_d} \right) + \frac{1}{X_d}] \tag{3-10}$$

For small variations $\Delta V, \Delta E, \Delta\alpha$:

$$\Delta P = \left.\frac{\partial P}{\partial V}\right|_{V_o} . \Delta V + \left.\frac{\partial P}{\partial E}\right|_{E_o} . \Delta E + \left.\frac{\partial P}{\partial \alpha}\right|_{\alpha_o} . \Delta\alpha \tag{3-11}$$

with $\dfrac{\partial P}{\partial \alpha} = \dfrac{V \cdot E \cdot \cos\alpha}{X_d} + V^2 \cos 2\alpha \left(\dfrac{1}{X_q} - \dfrac{1}{X_d} \right)$

$$\frac{\partial P}{\partial E} = \frac{V \cdot \sin\alpha}{X_d}$$

$$\frac{\partial P}{\partial V} = \frac{E \cdot \sin\alpha}{X_d} + V . \sin 2\alpha \left(\frac{1}{X_q} - \frac{1}{X_d} \right)$$

where V, E, α on the right hand side are steady state values.

Equations for generator (i)

- electrical torque

  with the appropriate substitutions:

$$\Delta T_{ei} \approx \Delta P_{ei} = K_{1i} \cdot \Delta\alpha_i + K_{2i} \cdot \Delta E'_q + K_{7i} \cdot \Delta V_L \tag{3-12}$$

with $K_{1i} \mathrel{\hat{=}} \dfrac{\partial P_{ei}}{\partial \alpha} = \dfrac{V_L . E'_{qi} \cdot \cos\alpha_i}{X'_{di} + X_t} + V_L^{\ 2} . \cos 2\alpha_i \left(\dfrac{1}{X_{qi} + X_t} - \dfrac{1}{X'_{di} + X_t} \right)$

$$K_{2i} \mathrel{\hat{=}} \frac{\partial P_{ei}}{\partial E'_{qi}} = \frac{V_L . \sin\alpha_i}{X'_d + X_t}$$

$$K_{7i} \hat{=} \frac{\partial P_{ei}}{\partial V_L} = \frac{E'_{qi}.\sin\alpha_i}{X'_d+X_t} + V_L.\sin 2\alpha_i \left(\frac{1}{X_{qi}+X_t} - \frac{1}{X'_{di}+X_t} \right)$$

K_{1i} and K_{2i} are equal to K_1 and K_2 in the paper of De Mello and Concordia [2].

- internal voltage Eq.:

$$\Delta E_{qi}=\Delta E'_{qi} + (X_{di}-X'_{di}).\Delta I_{di} = \Delta E'_{qi} + \frac{X_{di}-X'_{di}}{X'_{di}+X_t} (\Delta E'_{qi} -\Delta(V_L.\cos\alpha i))$$

or $\Delta E_{qi} = \frac{1}{K_{3i}} \cdot \Delta E'_{qi} + K_{4i}.\Delta\alpha_i + K_{8i}\cdot\Delta V_L$ (3-13)

with

$$K_{3i} = \frac{X'_{di}+X_t}{X_{di}+X_t}$$

$$K_{4i} = \frac{X_{di}-X'_{di}}{X'_{di}+X_t} \cdot V_L.\sin\alpha_i$$

$$K_{8i} = -\frac{X_{di}-X'_{di}}{X'_{di}+X_t} \cdot \cos\alpha_i$$

- terminal voltage V_t

$$V_t^2 = V_{dt}^2 + V_{qt}^2 \qquad \Delta V_t = \frac{\Delta V_{dt}\cdot V_{dt}+\Delta V_{qt}.V_{qt}}{V_t}$$

$$\Delta V_{dt}= X_{qi}\Delta I_{qi} = \frac{X_{qi}}{X_{qi}+X_t} \cdot \Delta(V_L\cdot\sin\alpha)$$

$$\Delta V_{qt}=\Delta E'_{qi} - X'_{di}.\Delta I_{di} = \frac{X_t}{X'_{di}+X_t}\cdot\Delta E'_{qi} + \frac{X'_{di}}{X'_{di}+X_t}\cdot\Delta(V_L.\cos\alpha)$$

hence $\Delta V_t= K_{5i}.\Delta\alpha_i+ K_{6i}.\Delta E'_{qi} + K_{9i}.\Delta V_L$ (3-14)

with

$$K_{5i}= \frac{V_{dt}}{V_t} \cdot \frac{X_{qi}}{X_{qi}+X_t}\cdot V_L\cdot\cos\alpha_i - \frac{V_{qt}}{V_t} \cdot \frac{X'_{di}}{X'_{di}+X_t} \cdot V_L.\sin\alpha_i$$

$$K_{6i} = \frac{V_{qt}}{V_t} \cdot \frac{X_t}{X'_{di}+X_t}$$

$$K_{9i} = \frac{V_{dt}}{V_t} \cdot \frac{X_{qi}}{X_{qi}+X_t} \cdot \sin\alpha_i + \frac{V_{qt}}{V_t} \cdot \frac{X'_{di}}{X'_{di}+X_t} \cdot \cos\alpha_i$$

- <u>terminal power</u>

$$\Delta P_t = \Delta P_L = \Delta T_{ei} \tag{3-15}$$

Again, the constants K_{3i} through K_{6i} are identical with K_3 through K_6 in [2].

- <u>voltage regulator</u>

$$T_{FD} \cdot \Delta\dot{E}_{FDi} + \Delta E_{FDi} = K_{FD}\,(\Delta V_{ref} - \Delta V_t) \tag{3-16}$$

- <u>governor and turbine</u>

$$T_G \cdot \Delta\dot{T}_{mi} + \Delta T_{mi} = K_G \cdot (\Delta\omega_{ref} - \Delta\omega_i) \tag{3-17}$$

ΔV_{ref} and $\Delta\omega_{ref}$ are input signals through which an intentional disturbance can be applied.

<u>Summary of equations for machine (i)</u>

$$\Delta\dot{\delta}_i = \omega_o \cdot \Delta\omega_i$$

$$M_i \cdot \Delta\dot{\omega}_i = T_{mi} - T_{ei} - D \cdot \Delta\omega_i$$

$$T_G \cdot \Delta\dot{T}_{mi} = -\Delta T_{mi} + K_G(\Delta\omega_{ref} - \Delta\omega_i)$$

$$T'_{doi} \cdot \Delta\dot{E}'_q = -\Delta E_{qi} + \Delta E_{FDi}$$

$$T_{FD} \cdot \Delta\dot{E}_{FDi} = -\Delta E_{FDi} + K_{FD}(\Delta V_{ref} - \Delta V_t)$$

where

$$\Delta T_{ei} = K_{1i}\cdot\Delta\alpha_i + K_{2i}\cdot\Delta E'_{qi} + K_{7i}\cdot\Delta V_L$$

$$\Delta E_{qi} = \frac{1}{K_{3i}}\cdot\Delta E'_{qi} + K_{4i}\cdot\Delta\alpha_i + K_{8i}\cdot\Delta V_L$$

$$\Delta V_t = K_{5i}\cdot\Delta\alpha_i + K_{6i}\cdot\Delta E'_{qi} + K_{9i}\cdot\Delta V_L$$

Summary of equations for machine (j)

$$\Delta\dot{\delta}_j = \omega_o \cdot \Delta\omega_j$$

$$M_j\cdot\Delta\dot{\omega}_j = -D_j\cdot\Delta\omega_j - \Delta T_{ej}$$

$$T'_{doj}\cdot\Delta\dot{E}'_{qj} = -\Delta E_{qj}$$

$$\Delta T_{ej} = K_{1j}\cdot\Delta\alpha_j + K_{2j}\cdot\Delta E'_{qj} + K_{7j}\cdot\Delta V_L$$

$$\Delta E_{qj} = \frac{1}{K_{3j}}\Delta E'_{qj} + K_{4j}\cdot\Delta\alpha_j + K_{8j}\cdot\Delta V_L$$

In the combined model for both machines, one machine angle can be eliminated, since only relative positions of the machine axes are of interest, $\Delta\delta \hat{=} \Delta\delta_i - \Delta\delta_j$. This reduces the order of the model by one.

3.4 Load Bus Equations

Three more equations are required to eliminate $\Delta V_L, \Delta\alpha_1, \Delta\alpha_2$ from the state equations. Two relations are from power continuity conditions:

$$P_L = P_{Lj} + P_{Li} \qquad\qquad Q_L = Q_{Lj} + Q_{Li}$$

and one is an angular relationship

$$\Delta\delta_1 - \Delta\delta_2 = \Delta\alpha_1 - \Delta\alpha_2 = \Delta\delta$$

hence

$$\left|\begin{array}{l} \Delta P_L = K_{1i}\Delta\alpha_i + K_{1j}\Delta\alpha_j + K_{2i}\Delta E'_{qi} + K_{2j}\Delta E'_{qj} + (K_{7i}+K_{7j})\cdot\Delta V_L \\ \\ \Delta Q_L = \quad \cdot \cdot \cdot \cdot \cdot \\ \\ \Delta\delta \;= \Delta\alpha_1 - \Delta\alpha_2 \end{array}\right| \qquad (3\text{-}18)$$

The 3 equations can be solved for $\Delta V_L, \Delta\alpha_i, \Delta\alpha_j$ by Gaussian elimination. This is best done numerically. The expressions are substituted in the state equations of the machines, which in effect will modify the constants $K_{1i} \ldots K_{6i}$.

3.5 Simplified Dynamic Model

Since the estimation algorithm posed some problems and is expensive in computing time, a simplified model was used:

- the voltage change on the load bus ΔV_L is assumed to be negligible
- the load bus is assumed to keep the frequency of the external machine

  $\Delta\alpha_i \approx \Delta\delta \quad \Delta\alpha_2 \approx 0$

Under these assumptions, only the equations of motion of the machine (j) remain relevant and its electrical torque can be expressed as

$$\Delta T_{ej} = \Delta P_L - \Delta T_{ei} \qquad (3\text{-}19)$$

State equations for simplified model:

$$\Delta\dot{\delta} = \omega_o(\Delta\omega_i - \Delta\omega_j) \qquad (3\text{-}20)$$

$$M_i \cdot \Delta\dot{\omega}_i = -D_i \cdot \Delta\omega_i + \Delta T_{mi} - \Delta T_{ei} \qquad (3\text{-}21)$$

$$M_j \cdot \Delta\dot{\omega}_j = -D_j \cdot \Delta\omega_j - \Delta T_{ej} \qquad (3\text{-}22)$$

$$T_G \cdot \Delta\dot{T}_{mi} = -\Delta T_{mi} + K_G(\Delta\omega_{ref} - \Delta\omega_i) \tag{3-23}$$

$$T'_{do} \cdot \Delta\dot{E}'_{qi} = \Delta E_{FD} - \Delta E_{qi} \tag{3-24}$$

$$T_{FD} \cdot \Delta\dot{E}_{FDi} = -\Delta E_{FDi} + K_{FD}(\Delta V_{ref} - \Delta V_t) \tag{3-25}$$

$$\Delta T_{ei} = K_{1i} \cdot \Delta\delta + K_{2i} \cdot \Delta E'_{qi} \tag{3-26}$$

$$\Delta E_{qi} = K_{4i}\Delta\delta + \frac{1}{K_{3i}} \Delta E'_{qi} \tag{3-27}$$

$$\Delta V_t = K_{5i}\Delta\delta + K_{6i}\Delta E'_{qi} \tag{3-28}$$

$$\Delta T_{ej} = \Delta P_L - \Delta T_{ei} \tag{3-29}$$

This is a 6th order model with the 3 unknown parameters M_j, D_j, X_t and the unknown covariance of the random input ΔP_L. Measurable quantities for the estimation procedure are: speed $\Delta\omega_i$, terminal voltage ΔV_t, terminal power $\Delta P_t = \Delta T_{ei}$.

3.6 Steady State Values

The steady state is computed from the known system parameters (M_i, X_{di}, ...), estimates of the unknown parameter X_t and the terminal steady state quantities V_{t0}, P_{t0}, Q_{t0}. The index 0 is omitted in the following.

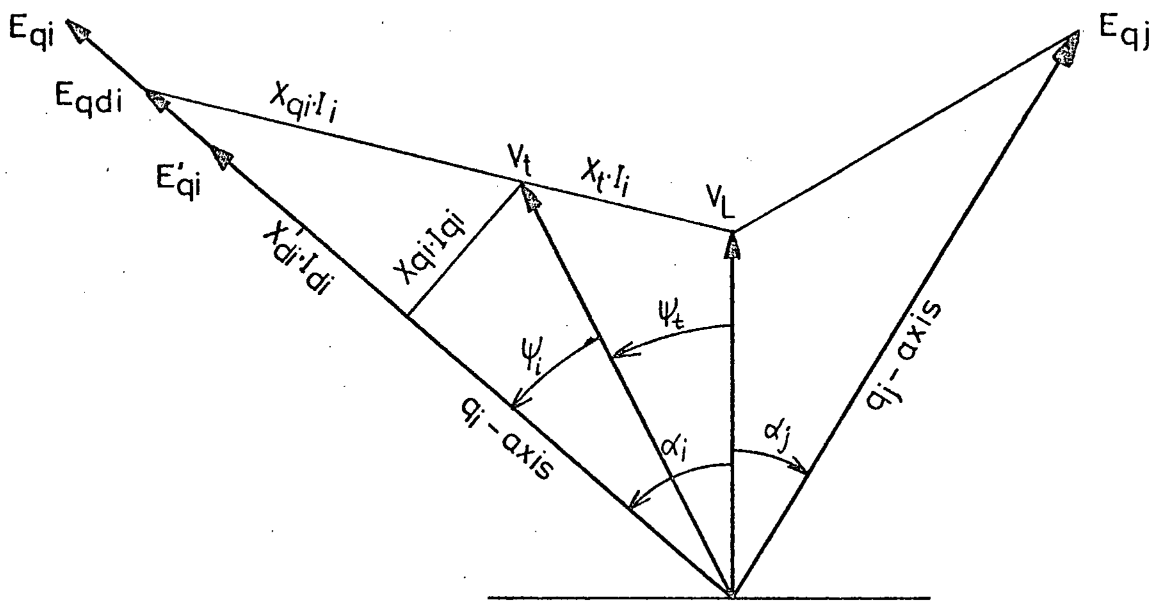


Fig. 3.3 Steady state phasor diagram

The equations (3-9, 3-10) applied for the terminal of machine (i) are:

$$P_t = \frac{V_t \cdot E_{qdi} \cdot \sin\psi_i}{X_{qi}} = \frac{V_t \cdot E'_{qi} \cdot \sin\psi_i}{X'_{di}} + V_t^2 \cdot \sin\psi_i \cos\psi_i \left(\frac{1}{X_{qi}} - \frac{1}{X'_{di}} \right)$$

$$Q_t = \frac{V_t \cdot E_{qdi} \cdot \cos\psi_i}{X_{qi}} - \frac{V_t^2}{X_q}$$

and

$$-P_t = \frac{V_t \cdot V_L \cdot \sin(-\psi_t)}{X_t} \qquad -Q_t = \frac{V_t \cdot V_L \cdot \cos(-\psi_t)}{X_t} - \frac{V_t^2}{X_t}$$

with the solutions

$$\psi_i = \text{arc tan} \frac{P_t}{Q_t + \frac{V_t^2}{X_{qi}}} \quad (3\text{-}29) \qquad \psi_t = \text{arc tan} \frac{P_t}{-Q_t + \frac{V_t^2}{X_t}} \quad (3\text{-}30)$$

$$E'_{qi} = \frac{P_t \cdot X'_{dt}}{V_t \cdot \sin\psi_i} + V_t \cdot \cos\psi_i \left(1 - \frac{X'_{di}}{X_{qi}} \right) \quad (3\text{-}31)$$

$$V_L = \frac{P_t \cdot X_t}{V_t \cdot \sin\psi_t} \quad (3\text{-}32) \qquad \alpha_i = \psi_i + \psi_t \quad (3\text{-}33)$$

4. IMPLEMENTATION OF ALGORITHM

This chapter contains a short description of the computer implementation of the estimator and of some methods to reduce computing time.

4-1 Calculation of Likelihood Values

The following flow-chart shows the major steps in the calculation of one likelihood value:

arguments α, Q, R

↓

calculation of system matrices $A(\alpha)$, $B(\alpha)$, $C(\alpha)$
equations of power system model, sections 3-5, 3-6

↓

conversion to discrete-time system $A(\alpha), B(\alpha) \rightarrow F(\alpha), G(\alpha)$
$C.Q.C^T \rightarrow Q_D = T \cdot F$
method see section 4-3

↓

calculation of steady-state Kalman gains K
and (expected) innovations covariance P_e
iterative equations, see section 2-3

↓

processing of the measurement arrays u, z
with the Kalman filter and summation of
the innovations (sample) covariance S
equations, see section 2-3

↓

likelihood value $L(\alpha, Q, R)$

The first of the steps above contains the physical model. The following steps are independent of it. The computation sequence is simpler for the case without input noise when there is no Kalman filter.

4-2 Optimization Procedure

The Monitor for nonlinear optimization available at UBC [14] was employed to find parameter estimates. It supports interactive processing and switching among different optimization methods. The following user routines were added:

FUNCTION XDFUNC - Calculation of a likelihood value as described in section 4-1

SUBROUTINE INIT - an initialization routine that must be called first and asks for the necessary data.

Most of the available algorithms in [14] were tried, but none of the methods based on gradient evaluation performed well. A possible cause is that the parameter steps taken for the difference approximation of the derivative were not suitable. Good results were achieved by the conjugate directions algorithm (routine POWEL).

4-3 Computation Efficiency

The number of function evaluations is high for any optimization algorithm. Since a function evaluation involves the setting up and running of a linear filter, computation time is long even for a small system. A considerable reduction is achieved by using the following techniques.

(a) Filter equations and matrix operations

The mathematical operations of the filter are simple but highly repetitive. Since multiplications of a vector by a matrix are among the most frequent operations, all matrices are stored by rows, so that $a_{i,j}$ becomes A(J,I) in Fortran. In most subroutines, array elements are addressed by a single index only.

A package of short matrix subroutines was written in Fortran. The order in which the elements are stored is chosen carefully. These matrix routines run several times faster than their more general counterparts in the system library.

(b) Conversion of continuous to discrete time

This can be done by numerical integration. If $x(0)$ is chosen as the i-th unit vector, then integration of the system $\dot{x}=Ax$ over one time step Δt will yield the i-th column of the transmission matrix F. The number of integration steps required is therefore equal to the order of the system. It was found that by using the Runge-Kutta method, up to 80% of the total computing time was spent on this integration.

A much faster procedure is obtained through approximation of the matrix exponential by its series:

$$F = \exp(A \cdot \Delta t) = I + \Delta t \cdot A + \frac{\Delta t^2}{2!} A^2 + \dots \qquad (4\text{-}1)$$

In general this method may suffer from several numerical problems, such as roundoff errors and slow convergence. But for the model in section 3-5, approximation by the first eight terms is accurate to 6 digits and about ten times faster than numerical integration.

The complete recursive algorithm for F, G and T is obtained from the series expansion of Eqs. (2-21, 2-22, 2-24):

$$F(n) = F(n-1) + \frac{1}{n!} \cdot \bar{A}(n) \qquad F(0)=I \qquad (4\text{-}2)$$

$$G(n) = G(n-1) + \frac{1}{n!} \cdot \bar{B}(n) \qquad G(0)=0 \qquad (4\text{-}3)$$

$$T(n) = T(n-1) + \frac{1}{n!} \cdot \bar{T}(n) \qquad T(0)=0 \qquad (4\text{-}4)$$

$$\bar{A}(n) = \bar{A}(n-1) \cdot \bar{A} \qquad \bar{A}(1)=\bar{A}=A\cdot\Delta t \qquad (4\text{-}5)$$

$$\bar{B}(n) = \bar{B}(n-1) \cdot \bar{B} \qquad \bar{B}(1)=\bar{B}=B\cdot\Delta t \qquad (4\text{-}6)$$

$$\bar{T}(n) = -\bar{T}(n-1)\cdot\bar{A}^{T}+\bar{A}(n-1)\cdot\bar{T} \qquad \bar{T}(1)=\bar{T}=C\cdot Q\cdot C^{T}\cdot\Delta t \qquad (4\text{-}7)$$

(c) Achieved computing times

The evaluation of a likelihood value for a system of 6th order, with one input, one output and 100 measurement points requires the following approximate CPU-times (for an Amdahl 470 computer):

24 msec for deterministic case (no Kalman gain)

55 msec for the stochastic case

A typical estimation run with 6 iterations of the conjugate directions algorithm requires about 100 function evaluations or 5 sec CPU-time. This is still considerable for such a simple model.

Further improvements are possible. Most of the calculations (especially the filter) could probably be done with single instead of double precision. Of most importance would be a faster, reliable method for computing the steady state Kalman gains.

4.4 Simulation Programmes

All data used for the estimation process was obtained from a computer simulation with the same model. Many of the subroutines were the same for estimation and simulation. The simulation data is therefore ideal and fulfills all assumptions of structural and noise properties that the estimator presupposes.

5 DATA AND RESULTS

The simulation data and the performance of the maximum likelihood estimation method are shown for the following cases:

Case 1 - deterministic input only

Case 2 - combined deterministic and stochastic inputs

Case 3 - stochastic input only

Case 4 - stochastic input and measurement noise

5.1 Parameters and Input Data

The following parameters were used in simulation program:

- sampling time step Δt=0.05 sec

(a) Deterministic parameters (all per unit)

- internal generator:

$M_i=5$	$D_i=10$	$T'_{di}=8$	
$X_{di}=1$	$X'_{di}=0.3$	$X_q=0.6$	
$K_{FD}=50$	$T_{FD}=0.1$	$K_G=10$	$T_G=0.5$

- external system (unknown parameters α)
  - transmission reactance: $X_t=0.5$
  - unknown machine: $M_j=9$ $D_j=26$

- steady state values:

$V_{t0}=1.05$ $P_{t0}=0.9$ $Q_{t0}=0.3$

(b) Deterministic input

The control input in case 1 and 2 is a short pulse applied at the reference input of the governor/turbine loop.

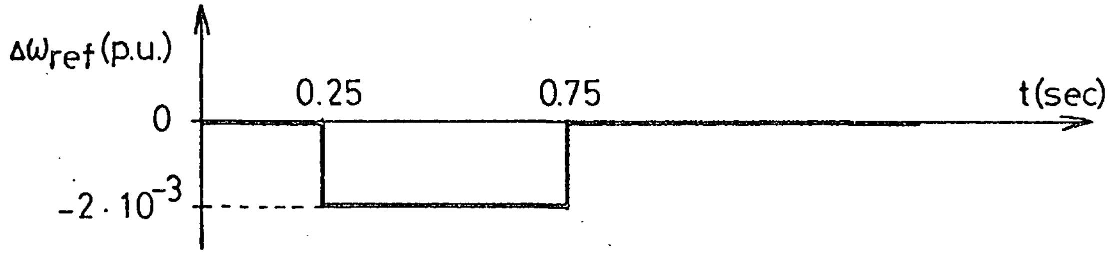


Fig. 5.1 Deterministic Input

(c) Stochastic input

A random disturbance ΔP_L was applied in cases 2 to 4. Only white noise with a variance $Q=E[\Delta P_L^2]=10^{-4}$ was considered. In discrete time this corresponds to a random sequence with a standard deviation of $\sigma_{PL}=\sqrt{Q\cdot\Delta t} = 2.2\cdot10^{-3}$ p.u.

For cases 1 to 3, measurement noise is neglected. In case 4, measurement noise with covariance $R=10^{-9}$ ($\sigma_v=0.032\cdot10^{-3}$) was added to the simulated output $\Delta\omega_i$.

5.2 Output Data

The simulated outputs were ΔP_t, ΔV_t and $\Delta\omega_i$. Each example consists of a sequence of 100 points. Some examples for each case are shown in Fig. 5.2 to 5.9. Table 5.1 contains for all simulated sequences the mean values m, the standard deviation σ, the maximum absolute amplitude $|A|$ and the expected innovation standard deviation $\sqrt{P_e}$ of a Kalman filter with the true parameters.

Table 5-1: Simulation Results:

Case	Signal	m		$\|A\|$	$\sqrt{P_e}$	Fig.
1	ΔP_t	-11.3	(23.9)	94.1	-	5.2
	ΔV_t	0.19	(2.02)	7.66	-	5.3
	$\Delta\omega_i$	-0.43	(0.62)	1.84	-	5.4
2	ΔP_t	-8.78	(30.3)	85.7	same	5.5
	ΔV_t	0.31	(4.10)	11.2	as	-
	$\Delta\omega_i$	-0.33	(0.83)	2.46	case 3	5.6
3	ΔP_t	2.53	21.8	42.7	3.33	5.7
	ΔV_t	0.120	3.07	6.26	0.388	-
	$\Delta\omega_i$	0.099	0.445	0.881	0.0268	5.8
4	$\Delta\omega_i$	0.1011	0.447	0.884	0.0806	5.9

All values x 10^{-3} per unit.

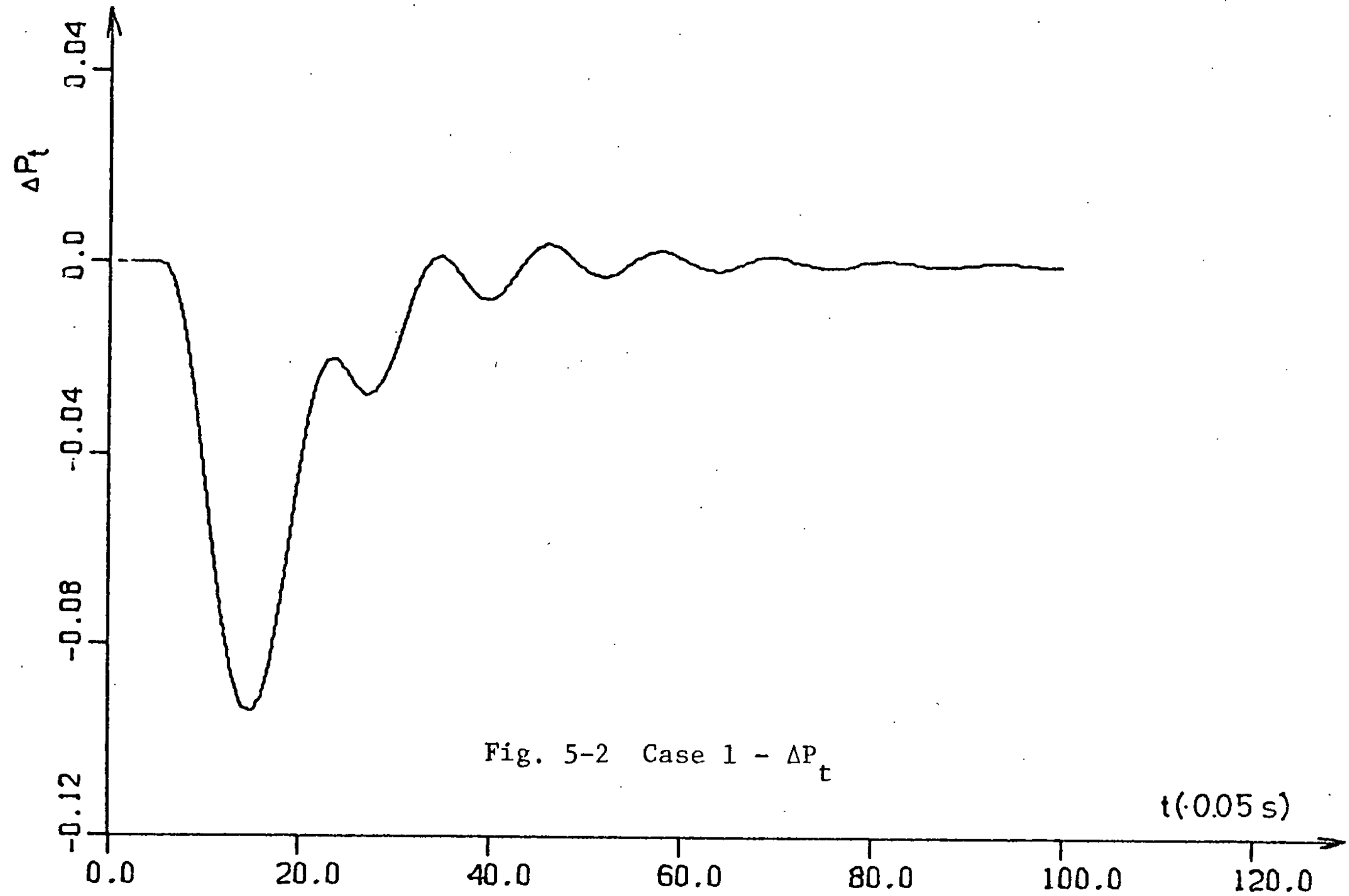


Fig. 5-2 Case 1 - ΔP_t

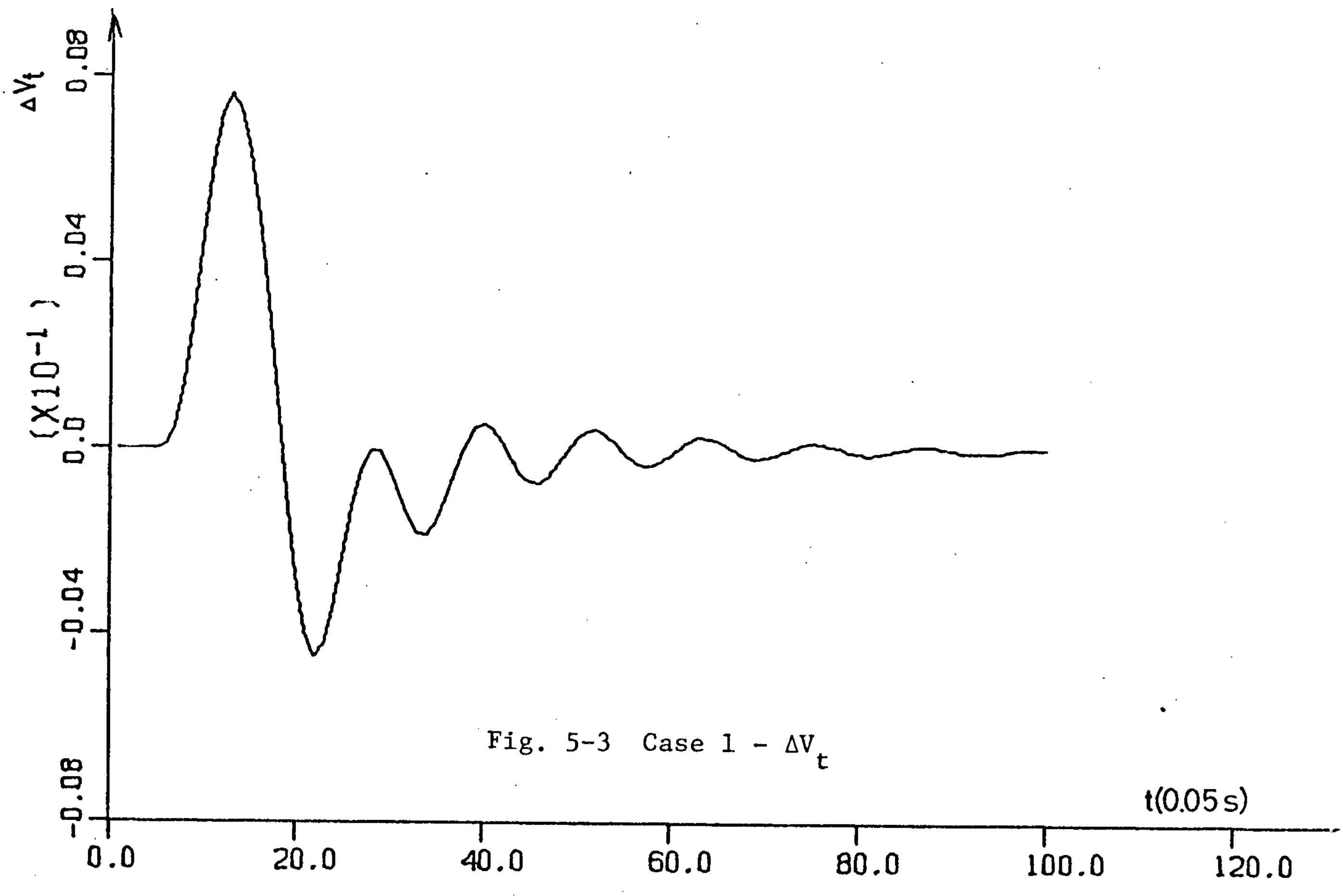


Fig. 5-3 Case 1 - ΔV_t

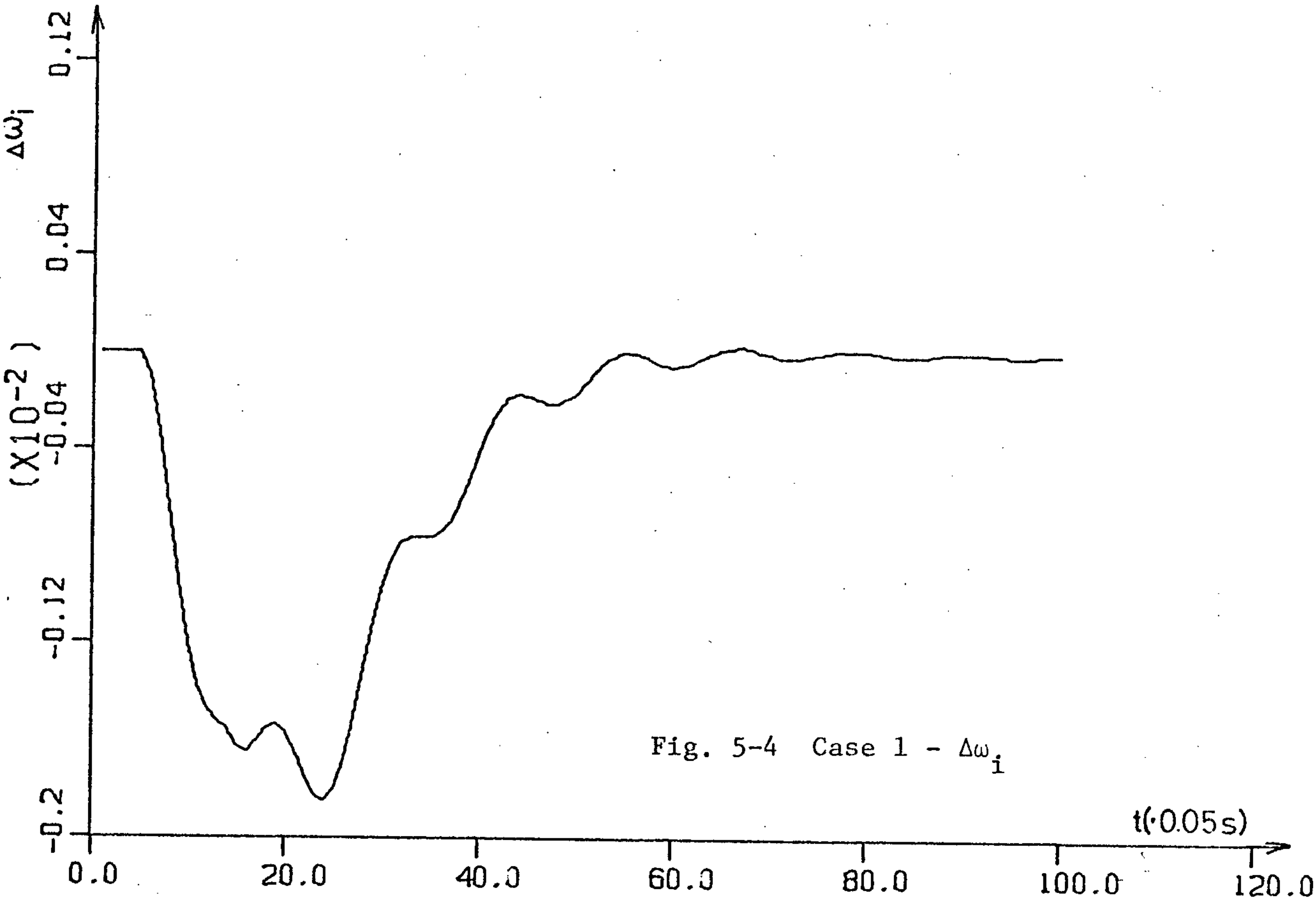


Fig. 5-4 Case 1 - $\Delta\omega_i$

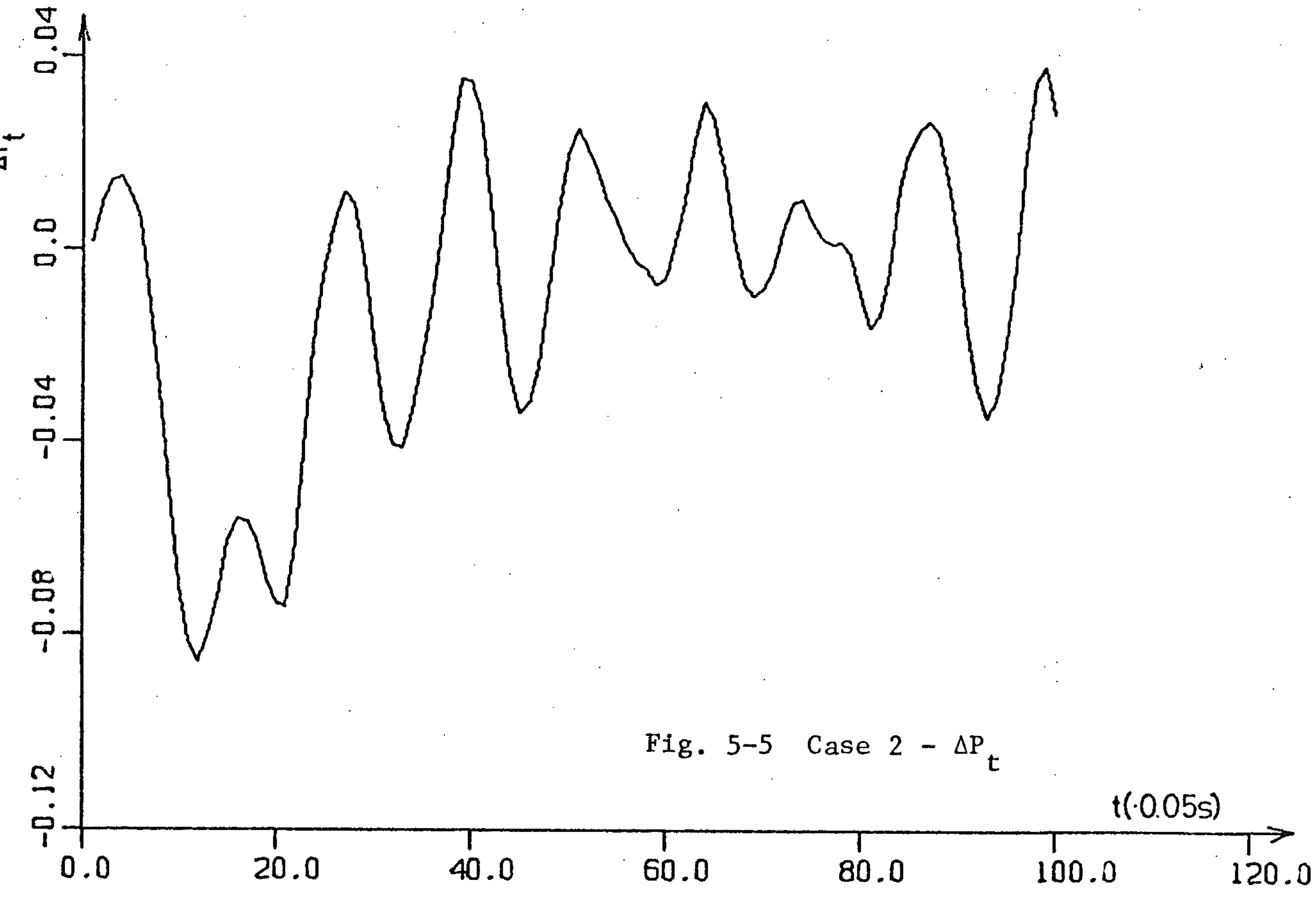


Fig. 5-5 Case 2 - ΔP_t

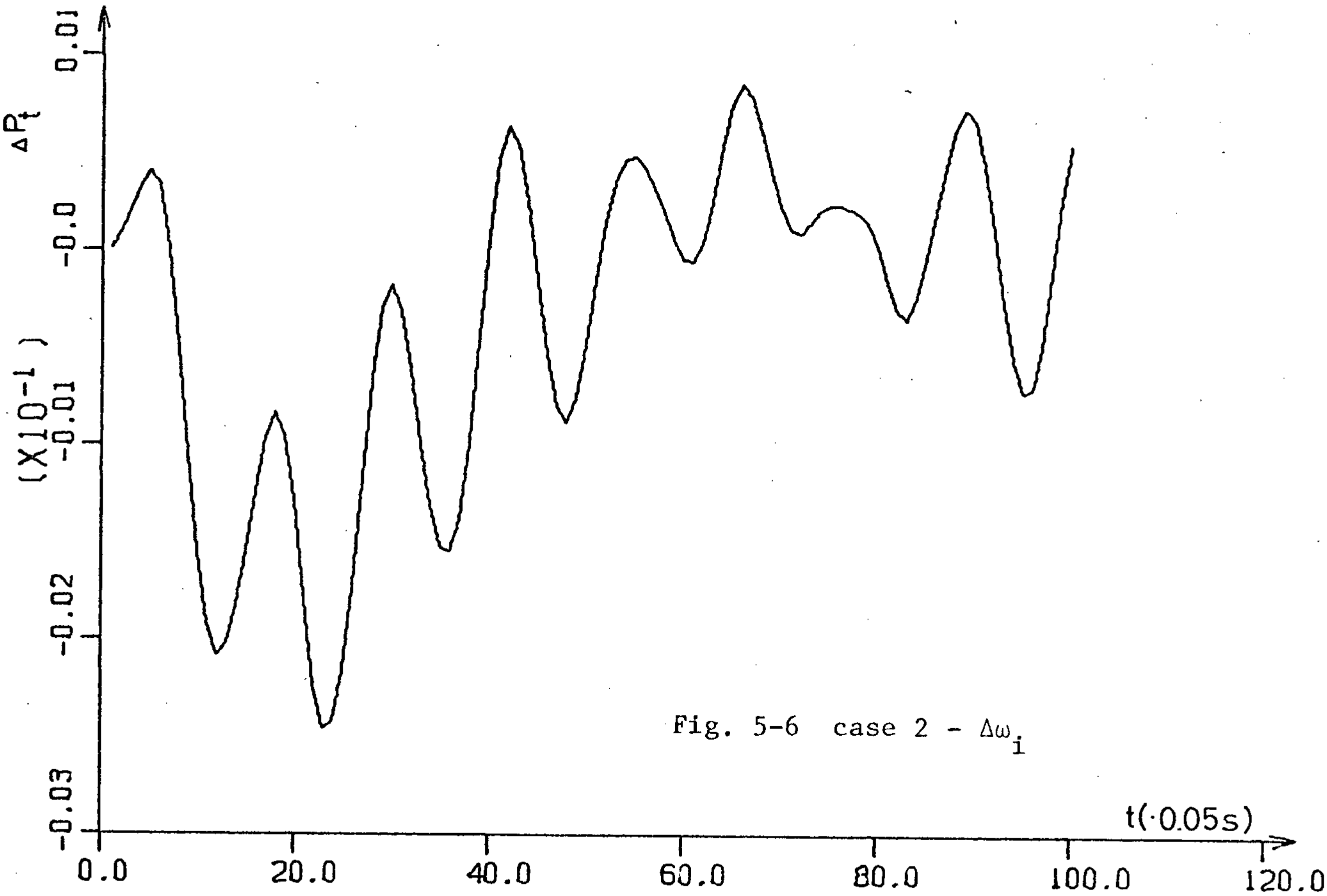


Fig. 5-6 case 2 - $\Delta\omega_i$

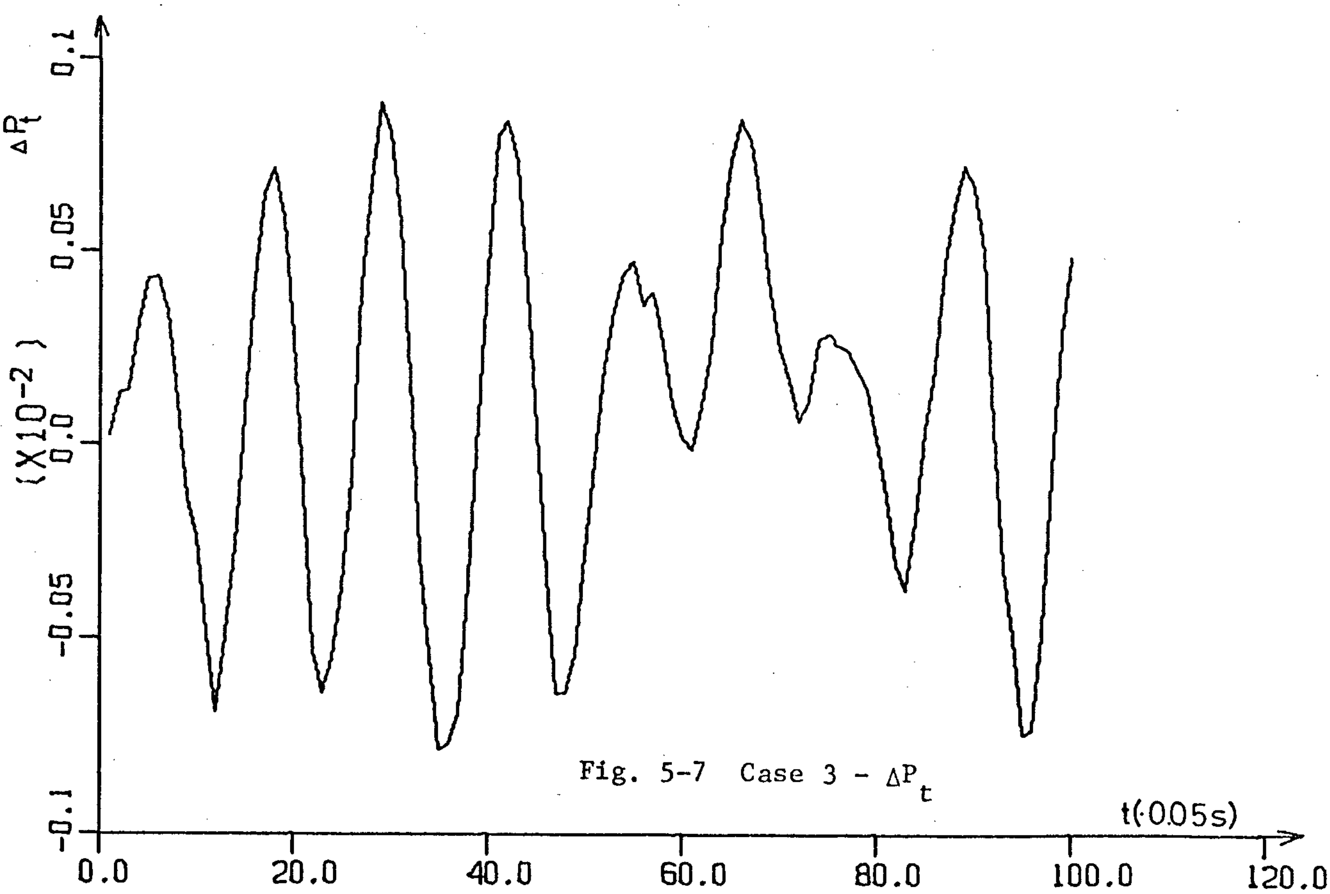


Fig. 5-7 Case 3 - ΔP_t

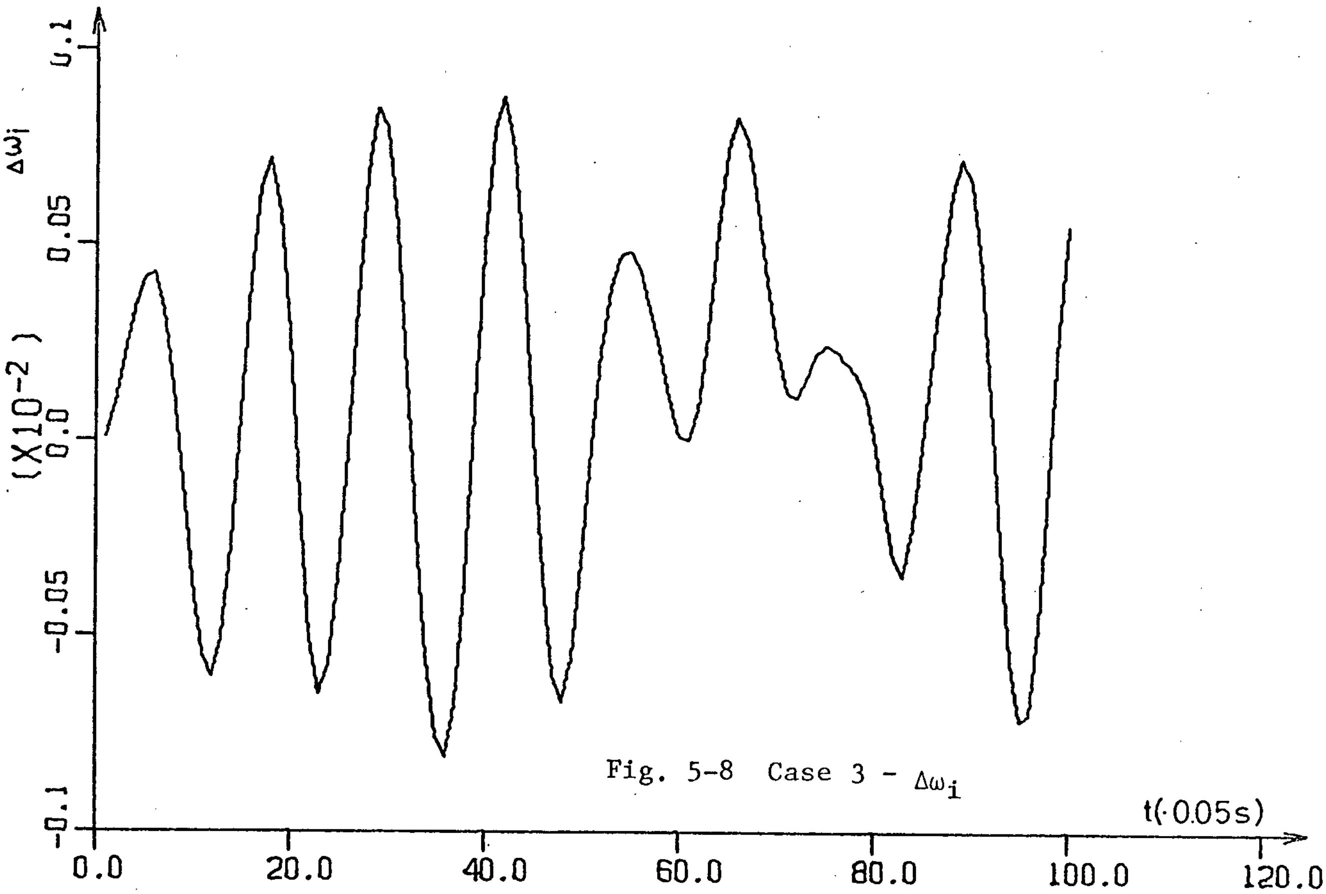


Fig. 5-8 Case 3 - $\Delta\omega_i$

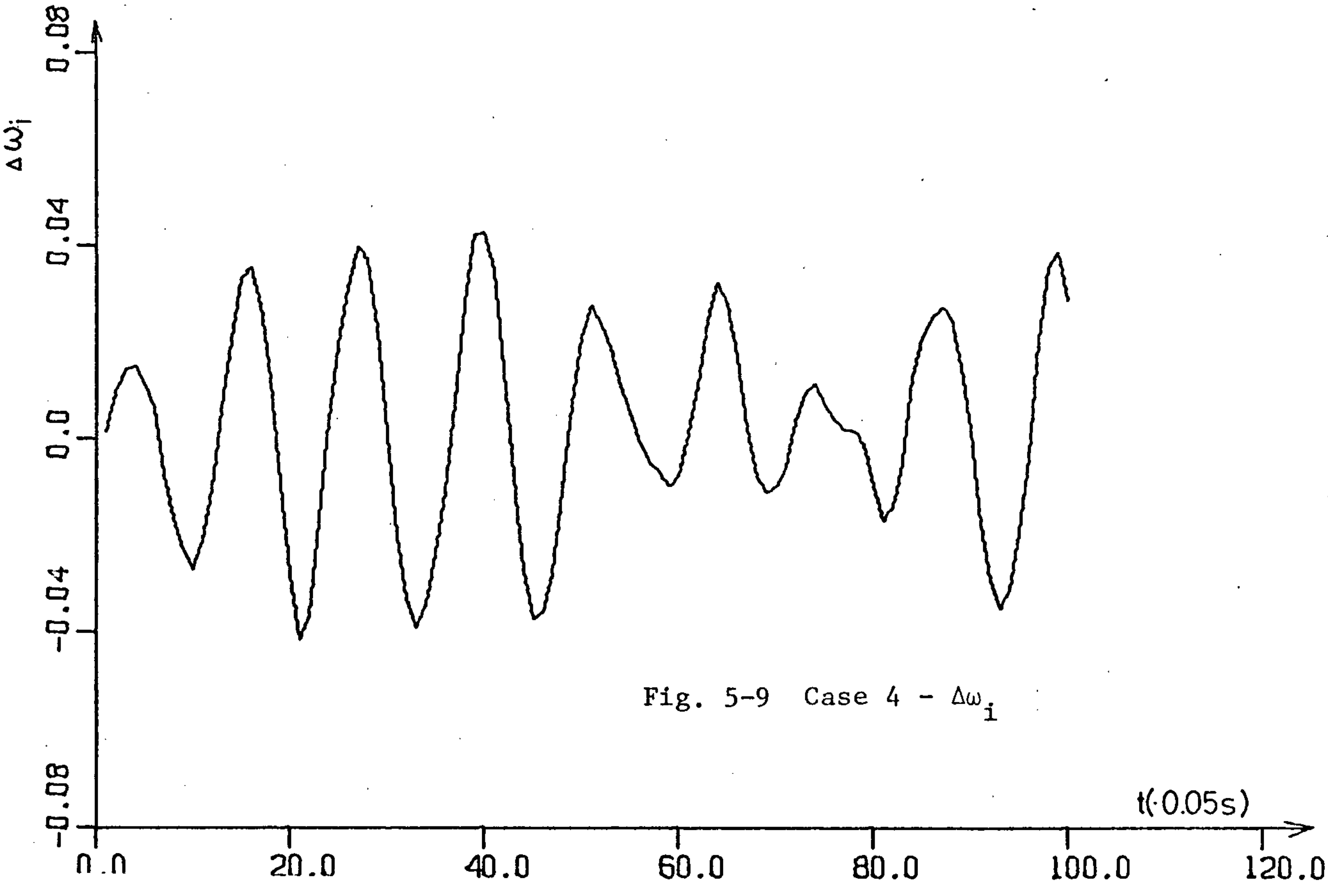


Fig. 5-9 Case 4 - $\Delta\omega_i$

5.3 Convergence of parameter estimate

Parameters were estimated from the simulated measurements for all four cases. In all examples the cost function was minimized by the conjugate directions algorithm (routine POWEL[14]). Only single output sequences were considered. This is sufficient in the cases without measurement noise, because one output contains all available information about the system. Estimation from multiple output sequences was not attempted, because numerical problems were encountered in the calculation of the Kalman gains. This should be further investigated for the case with measurement noise.

Table 5-2 contains a summary of the obtained optimal estimates for all examples. Fig. 5-10 to 5-18 show how the estimates converged and are followed by comments for the individual cases.

Table 5-2 Estimation Results

Case	Output	Initial guess $\hat{\alpha}_o$	# of iterations	Computing time [sec]	Optimal estimate $\hat{\alpha}^*, \hat{Q}^*$ and bias in % M_j	%	D_j	%	X_T	%	Q $[\times 10^{-4}]$	Fig.
1	ΔP_t	1)	6	2.15	9.000	-	25.999	-	50.00	-	-	5-10
2	ΔP_t	2)	5	6.0	8.25	-8.3	25.07	-3.6	51.08	+2.2	0.998	5-11
	$\Delta \omega_i$	2)	6	10.3	9.07	+0.8	27.7	+6.5	50.01	-	0.999	5-12
	$\Delta \omega_i$	3)	8	11.2	9.08	+0.9	27.7	+6.5	50.01	-	0.999	5-13
3	ΔP_t	2)	4	3.6	8.80	-2.2	30.0	+15.8	50.64	+1.3	1.005	5-14
	ΔV_t	2)	5	4.8	8.92	-0.9	31.7	+21.9	51.02	+2.0	1.012	5-15
	$\Delta \omega_i$	2)	4	4.8	8.72	-3.1	30.2	+16.2	50.8	+1.6	1.011	5-16
4	$\Delta \omega_i$	2)	4	5.8	8.64	-4.0	30.7	+18.1	50.1	+0.2	?	5-17
	$\Delta \omega_i$	4)	4*	15.5	9.48	+5.3	35.2	+35.	48.0	-4.0	1.172	5-18

1) α_o = (5., 5., 0.2)

2) α_o = (5., 5., 0.3)

3) α_o = (30, 10, 0.2)

4) α_o = (5, 5, 0.3), $Q_o = 10^{-6}$

*diverged later

N.B. the true parameters are α_N=(9., 26., 0.5), $Q=10^{-4}$, ($R=10^{-9}$ in case 4)

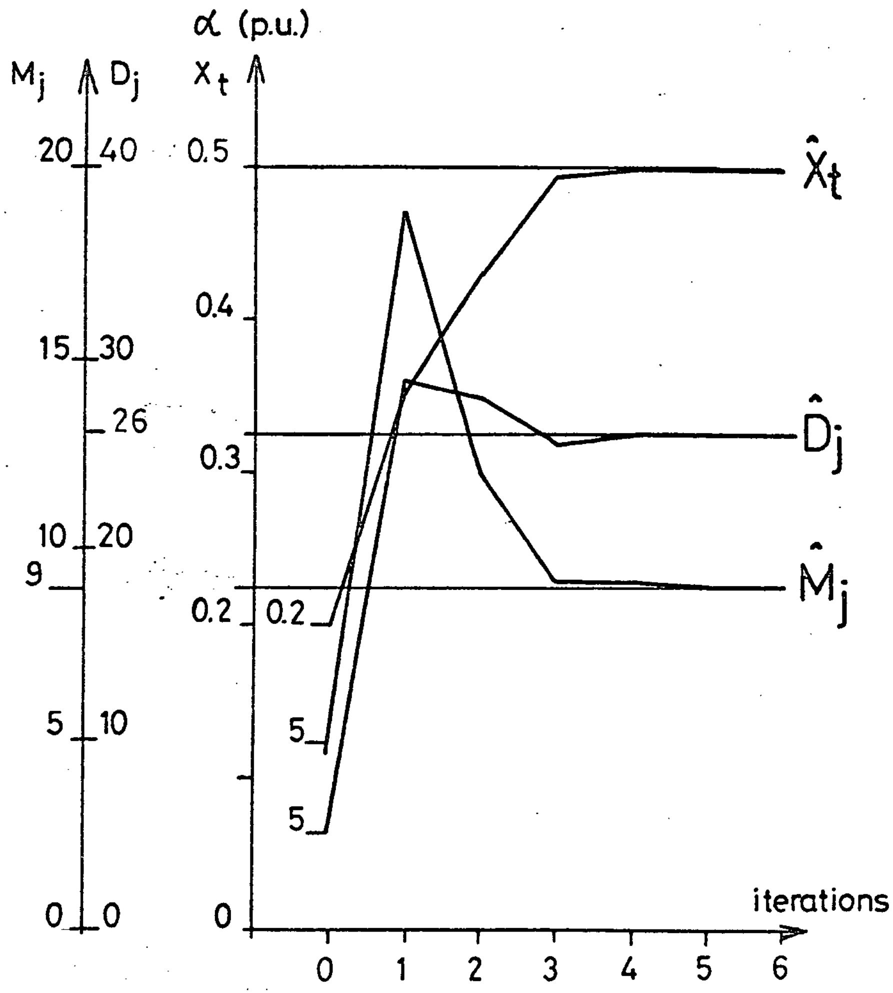


Fig. 5-10 Case 1 - deterministic output ΔP_t

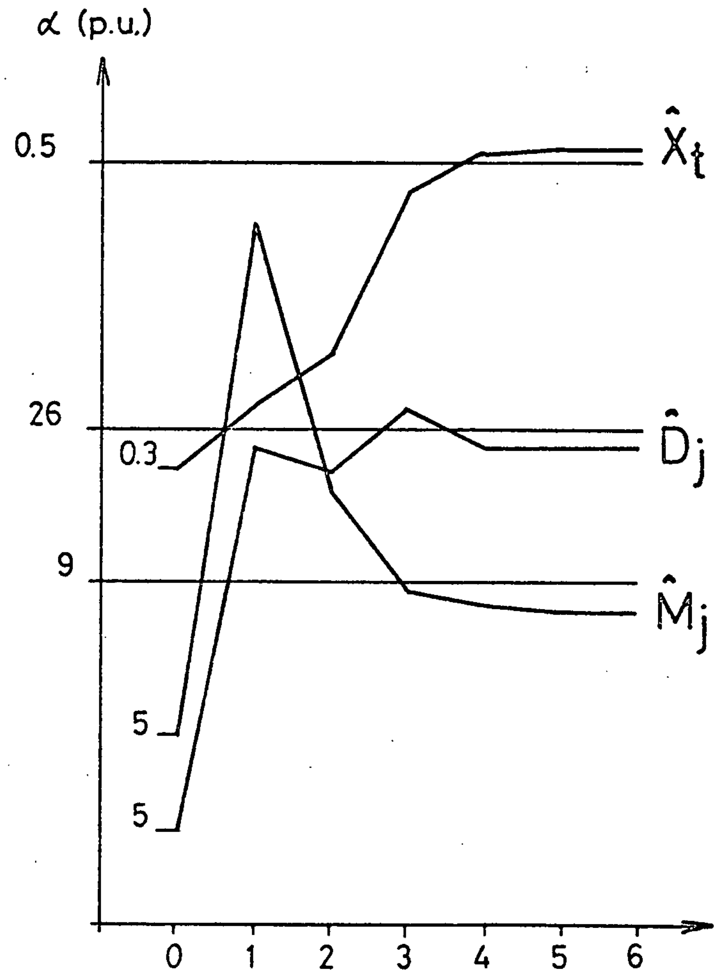


Fig. 5-11 Case 2 - mixed input output ΔP_t

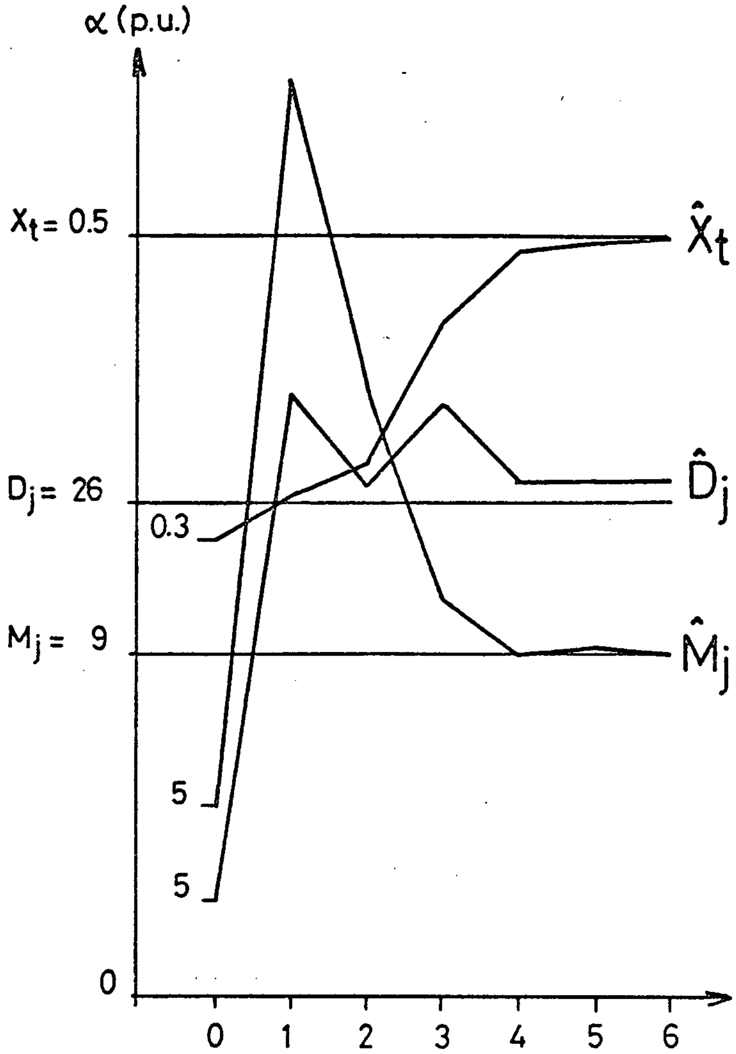


Fig. 5-12 Case 2 - mixed input

output $\Delta\omega_i$

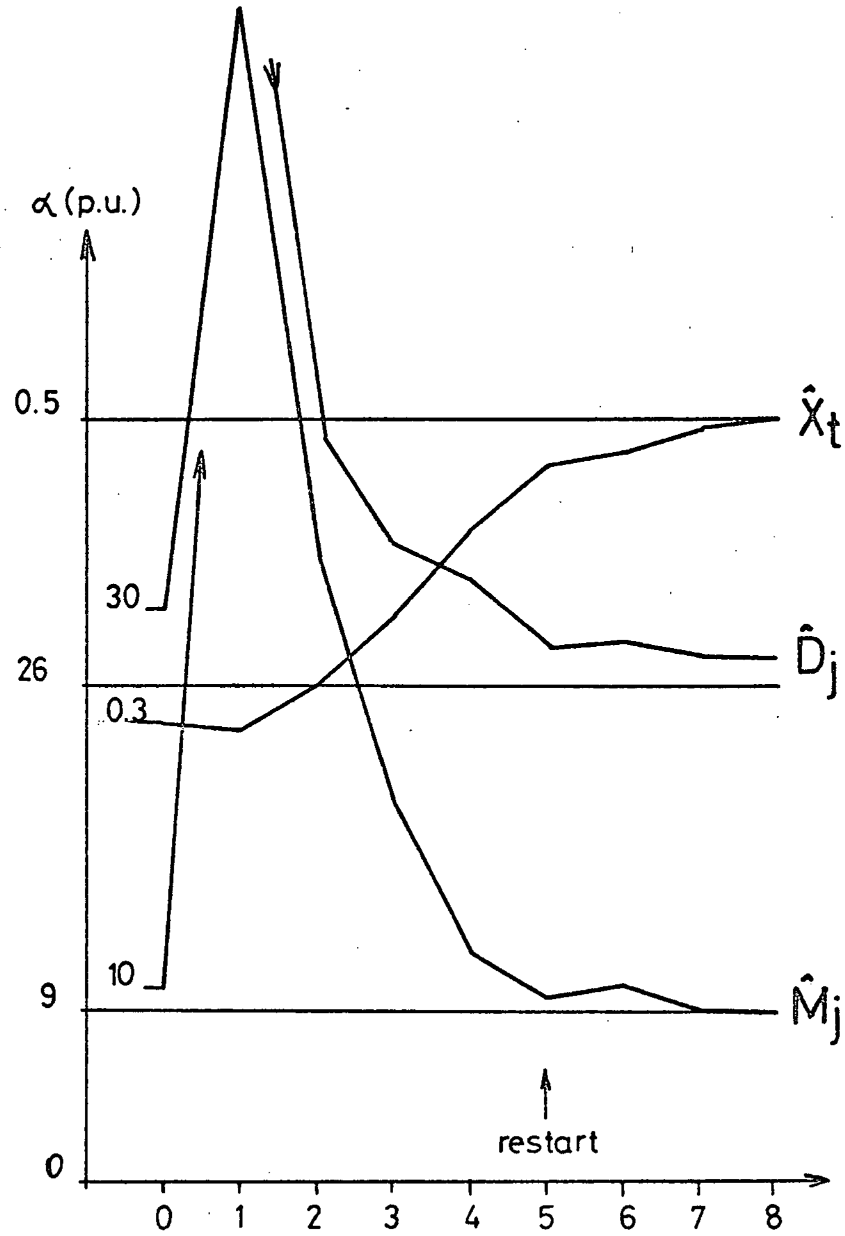


Fig. 5-13 Case 2 - mixed input

output $\Delta\omega_i$

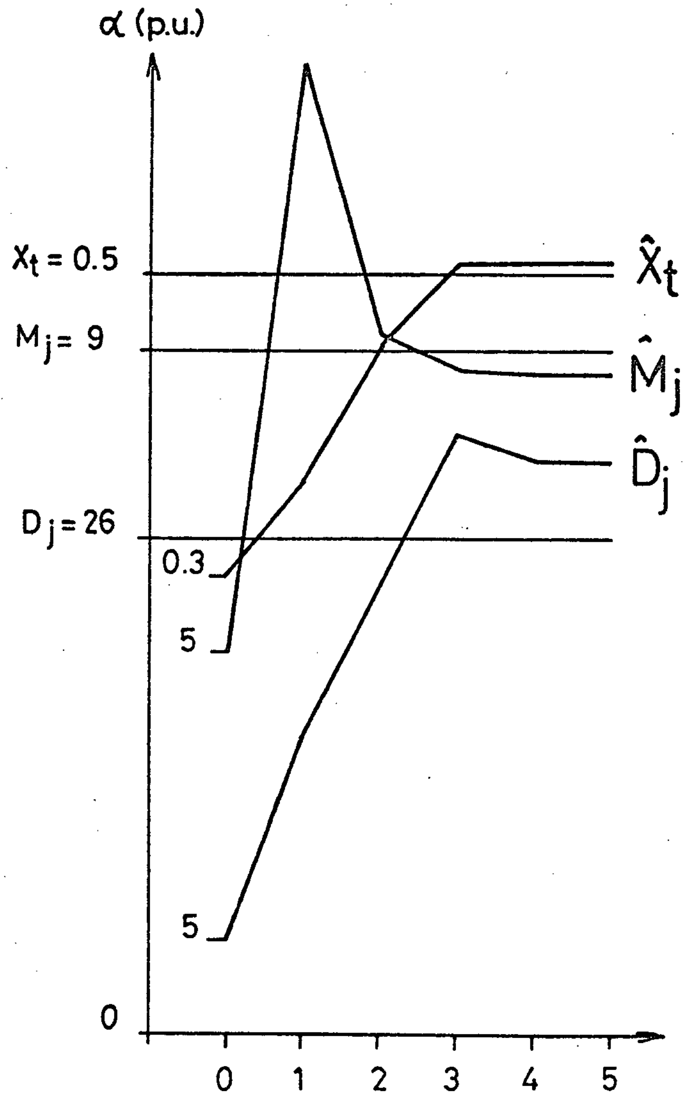


Fig. 5-14 Case 3 - stochastic output ΔP_t

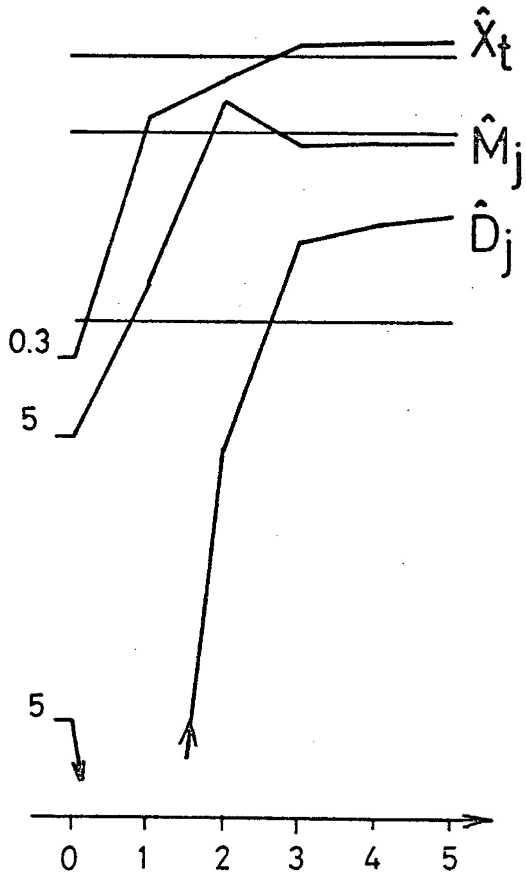


Fig. 5-15 Case 3 - stochastic output ΔV_t

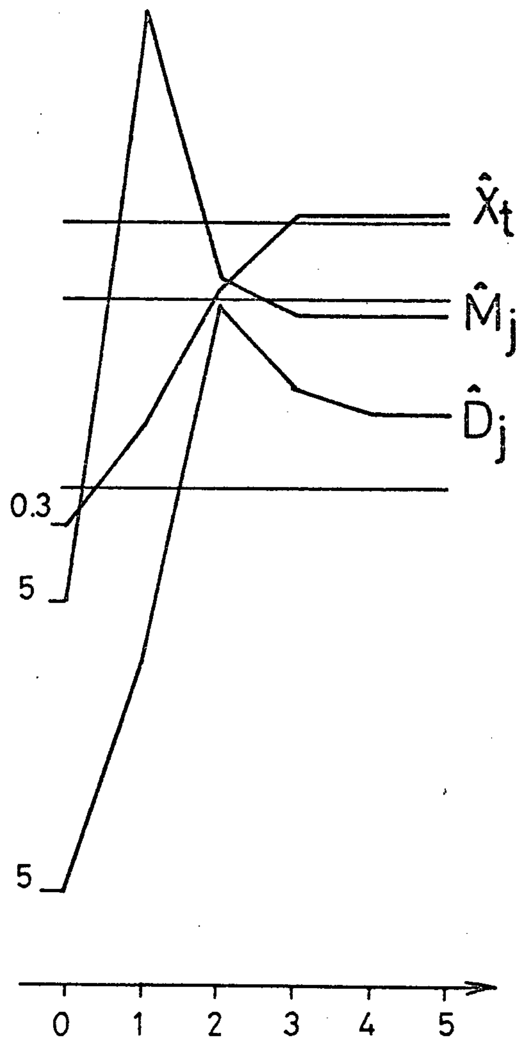


Fig. 5-16 Case 3 - stochastic output $\Delta\omega_i$

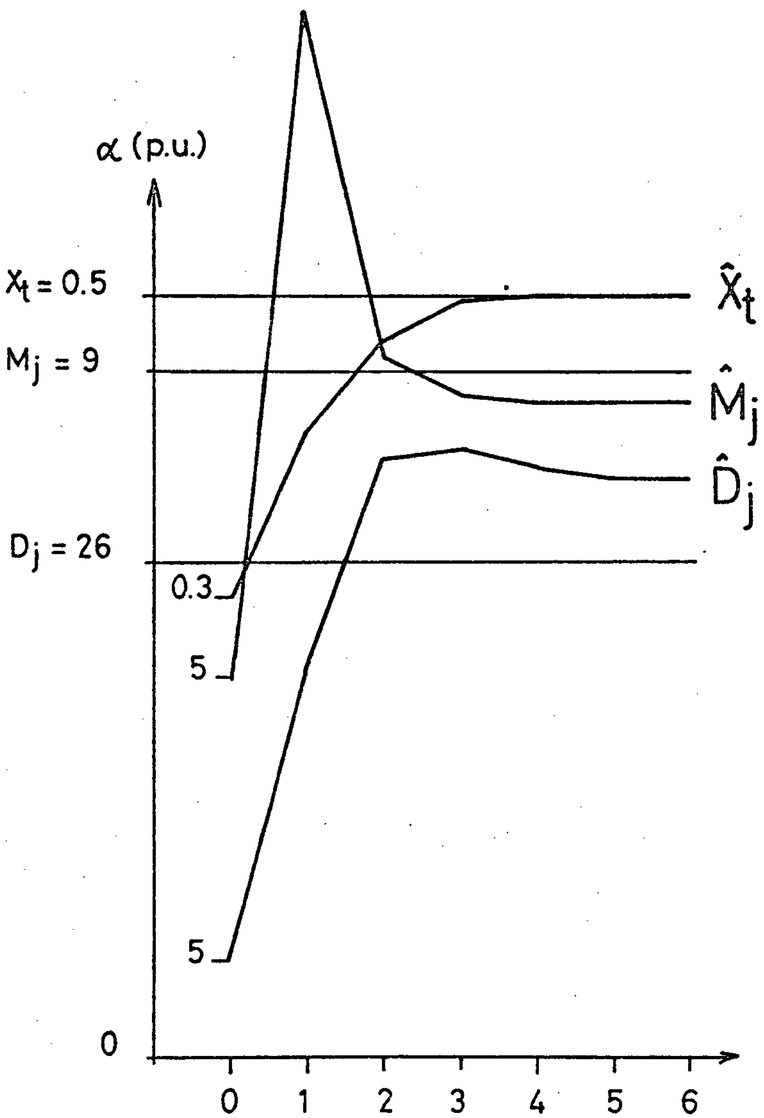


Fig. 5-17 Case 4 - with measurement noise output $\Delta\omega_i$

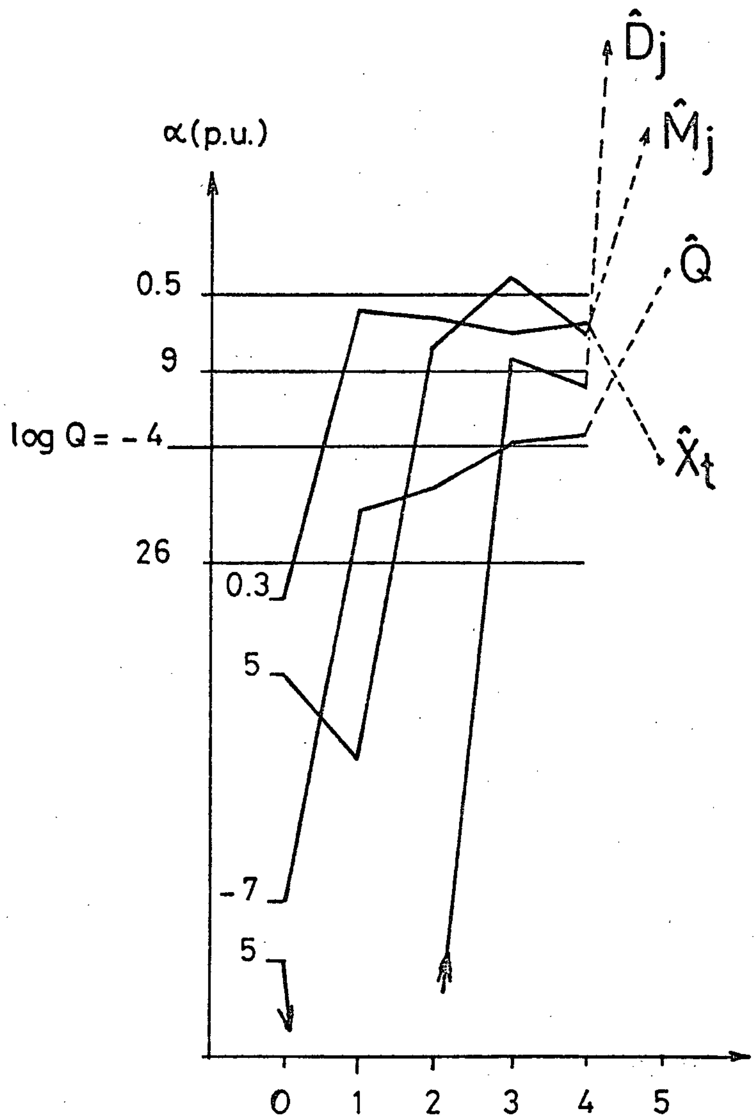


Fig. 5-18 Case 4 - with measurement noise output $\Delta\omega_i$

(a) Case 1 - deterministic

In this special case there is no Kalman filter. The estimator is the ordinary least squares estimator with the cost function

$$V = \sum_k e^2(k).$$

Rapid convergence to the true parameters was obtained for ΔP_t(Fig. 5-10) and $\Delta\omega_i$ measurements. The algorithm failed for the ΔV_t measurements, because the initial steps were too large.

(b) Case 2 - deterministic and stochastic input

The same cost function was used, but now the e(k) are the innovations of a Kalman filter. For negligible measurement noise ($R\to 0$), the Kalman gains depend on the parameters α only and the expected innovations covariance P_e is proportional to Q. Q is therefore estimated from the sample covariance $\hat{P}_e = V/N$.

Speed of convergence is comparable to the deterministic case, but the estimates have some bias (Fig. 5-11, 5-12). The first step of the algorithm always seems to make $\hat{M}_j$ and $\hat{D}_j$ larger, even if they are already too large (Fig. 5-13). It seems preferable to start with estimates that are smaller than expected.

(c) Case 3 - stochastic input only

Estimation with the cost function $V= \sum e^2(k)$ did not bring satisfactory results. Slightly biased estimates of M_j and D_j were found if X_t was held constant at the true value 0.5. But this is only one local minimum of V and lower minima can be found for quite different values of $\hat{M}_j$, $\hat{D}_j$ and $\hat{X}_t$.

Good results were obtained with the original likelihood cost function $L = \ln \det P_e + \sum e^2(k)/P_e$, where P_e is the estimated covariance of the Kalman filter. Estimates from all of the three outputs converged (Figs. 5-14 to 5-17), with a small bias for $\hat{M}_j$ and $\hat{X}_t$ and a rather large bias for $\hat{D}_j$.

(d) Case 4 - input and measurement noise

In the first example (Fig. 5-17) Q and R are held constant at the true values and only α is optimized. The results are similar as in case 3.

In the second example (Fig. 5-18) only R is fixed, α and Q are estimated simultaneously. After four iteration steps, the estimates seem to converge reasonably. However, in the next iteration step very different estimates were found, for slightly lower function value. For this short measurement sequence, the algorithm will not converge to the true values.

6 CONCLUSIONS

The maximum likelihood method is a general method for estimation of parameters in a linear system disturbed by both input and measurement noise. The stochastic parameters can be modelled either as unknown gains of a linear filter or as parameters of a low order model derived from physical knowledge.

The second approach is given preference and a low order model for the remote part of a power system is proposed that explains small fluctuations by varying load demand.

Estimation results were obtained for a simplified system from one output at a time. They show that the maximum likelihood method can in fact produce good estimates for a heavily disturbed system. Parameter estimates from a system with noise input only are also possible. However, convergence problems may arise for short measurement sequences. False convergence to a wrong set of parameters is possible, especially if the measurement noise is appreciable. This case would need further investigation with longer data sequences and a more realistic power system model.

Estimation of the dynamics of a real power system from measurements alone should be possible if

- accurate measurements of small fluctuations are available.
- the dynamic behaviour and the source of fluctuations of the external system with respect to a generator can be represented with sufficient accuracy by a very low order model.

To answer these questions, actual measurements from a generator should be investigated.

7 BIBLIOGRAPHY

1. Y.N. Yu and H.A. Moussa, "Optimal stabilization of a multi-machine system", IEEE Trans. on PAS, vol. PAS-91, pp. 1174-82 (May 1972).
2. F.P. DeMello and C. Concordia, "Concepts of synchronous machine stability as affected by excitation control", IEEE Trans. on PAS, vol. PAS-88, pp. 316-329 (April 1969).
3. Y.N. Yu, M.A. El-Sharkawi and M.D. Wvong, "Estimation of unknown large power system dynamics", IEEE Trans. on PAS, vol. PAS-98, pp. 279-289 (January 1979).
4. S. Lindahl and L. Ljung, "Estimation of power generator dynamics from normal operating data", 3rd IFAC Symposium, The Hague (June 1972), Part I, pp. 367-374, Paper pp 2.
5. W. Price, F.C. Schweppe, E.M. Gulachenski and R.F. Silva, "Maximum likelihood identification of power system dynamic equivalents", Proceedings 1974 IEEE Conference on Decision and Control, pp. 579-586.
6. R.D. Masiello and F.C. Schweppe, "Multi-machine excitation stabilization via system identification", IEEE Trans. on PAS, vol. PAS-94, pp. 444-454, (March 1975).
7. W. Price, D.N. Ewart, E.M. Gulachenski and R.F. Silva, "Dynamic equivalents from on-line measurements", IEEE Trans. on PAS, vol. PAS-94, pp. 1349-1357, (July 1975).
8. K.J. Åström and P. Eykhoff, "System identification - a survey", Automatica, vol. 7 (1971), pp. 123-161.
9. A. Gelb (ed.), "Applied optimal estimation", The Analytic Sciences Corporation (MIT-Press), 1974, 74-1604.

10. F.C. Schweppe, "Uncertain dynamic systems", Prentice Hall, 1973, 72-11801.

11. K.J. Åström, "Modelling and identification of power system components", from E. Handschin (ed.), "Real time control of electric power systems", Elsevier Publishing Company, 1972.

12. R.L. Kashyap, "Maximum likelihood identification of stochastic linear systems", IEEE Trans. on AC, vol. AC-15, pp. 25-34 (Feb. 1970).

13. P.R. Adby and M.A.H. Dempster, "Introduction to optimization methods", Chapman and Hall mathematic series, 1974, 74-4109.

14. M. Patterson, "UBC-NLP, Nonlinear function optimization", Computing Centre, Univ. of B.C., 1978.

15. Y. Bard, "Comparison of gradient methods for the solution of nonlinear parameter estimation problems", SIAM Journal Numer. Anal., vol. 7, pp. 157-186, (March 1970).

16. E.W. Kimbark, "Power system stability: Synchronous machines", Dover, 1968, 68-12937.