AN ABSTRACT OF THE THESIS OF

Talal Umar Halawani for the degree of Doctor of Philosophy

in Electrical and Computer Engineering presented on March 1, 1983

Title: AN OPTIMAL CONTROL APPROXIMATION FOR A CERTAIN CLASS OF NONLINEAR FILTERING PROBLEMS

Redacted for privacy

Abstract approved: ______________________

Ronald R. Mohler

A new approximation technique to a certain class of nonlinear filtering problems is considered in this dissertation. The method is based on an approximation of nonlinear, partially-observable systems by a stochastic control problem with fully observable state. The filter development proceeds from the assumption that the unobservables are conditionally Gaussian with respect to the observations initially. The concepts of both conditionally Gaussian processes and an optimal-control approach to filtering are utilized in the filter development. A two-step, nonlinear, recursive estimation procedure (TNF), compatible with the logical structure of the optimal mean-square estimator, generates a finite-dimensional, nonlinear filter with improved characteristics over most of the traditional methods. Moreover, a "close" (in the mean-square sense) approximation model for the original system will be generated as well. In general the nonlinear filtering problem does not have a finite-dimensional recursive synthesis. Thus, the proposed technique

may expand the range of practical problems that can be handled by nonlinear filtering. A detailed derivation for the filter with global property is presented. Extension of the results to large-scale nonlinear systems is accomplished by incorporating a novel decomposition scheme in the filter design.

Application of the developed filter to a scalar nonlinear system which lacks model "smoothness" is presented in [K2]. Application of the derived multi-dimensional filtering algorithm to two low-order, nonlinear tracking problems according to a global criterion and a local-time criterion respectively are presented. Also, a comparison with traditional methods, such as the popular Extended-Kalman Filter (EKF), are given via digital-computer simulation to demonstrate the effectiveness of the obtained results.

AN OPTIMAL CONTROL APPROXIMATION FOR A CERTAIN CLASS OF NONLINEAR FILTERING PROBLEMS

by

TALAL UMAR HALAWANI

A THESIS

submitted to

OREGON STATE UNIVERSITY

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

June, 1983

APPROVED:

Redacted for privacy

Professor of Electrical and Computer Engineering in Charge of Major

Redacted for privacy

Head of Department of Electrical and Computer Engineering

Redacted for privacy

Dean of Graduate School

Date thesis is presented March 1, 1983

Typed by WORD PROCESSING SPECIALISTS for Talal Umar Halawani

Acknowledgements

I would like to express my sincere gratitude to Professor Ronald R. Mohler, Chairman of my Dissertation Committee, for his valuable guidance, criticism, endless patience and encouragements throughout the course of this research.

I would like to express my sincere appreciation to Dr. W. K. Kolodziej whose assistance and invaluable suggestions made this dissertation possible. My gratitude also goes to the other members of my Dissertation Committee Professors J. K. Knudsen, R. Engelbrecht, and L. Murphy for their encouragement throughout this research.

My special thanks go to my wife Munira for the years of patience, understanding and encouragement.

This research was supported in part by the Office of Naval Research through Contract No. N00014-81-K-0814.

Last, but not least, I would like to thank the University of Petroleum and Minerals, Dhahran, Saudi Arabia, for the support which has made my graduate study possible.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

LIST OF FIGURES

LIST OF FIGURES
(Continued)

LIST OF FIGURES
(Continued)

LIST OF TABLES

NOTATIONS

The following notations will be used throughout this dissertation:

R^n ,	Euclidean n - dimensional space
C_T ,	Space of continuous functions on $[0,T]$
T ,	Time interval, usually $(-\infty, \infty)$
Ω ,	Sample space
F ,	Field
ζ_t ,	Sub σ-algebra $\zeta_t \leqslant F$
P ,	Probability measure
(Ω, F, P) ,	Complete probability space
tr (.) ,	Trace operator
E (.) ,	Expectation operator
$E(x/\ )=\bar{x}_t$,	Conditional expectation, $E(x(t)/y_s\ ;\ s\leqslant t)$
Γ_t ,	Conditional-error covariance, $E((x(t)-\bar{x}(t)(x(t)-\bar{x}(t)^*/y_s;\ s\leqslant t)$
$x\varepsilon[\ ]$,	An element of the subset [ ]
* ,	Transposition operator
$\|\ \|$,	Euclidean norm
$t\Lambda s$,	Min (s,t)
[ ] ,	Refer to references
(i-j) ,	Refer to equation or set of equations where i denotes the chapter, and j is the serial (sequential) number of that equation
$V_t = \frac{\partial V}{\partial t}$,	First-order derivative with respect to t (time)

$\frac{\partial V}{\partial z_i} = V_z$, i^{th} ,	First-order derivative (Gradient vector)
$\frac{\partial^2 V}{\partial z_i \partial z_j} = V_{zz}$, ij^{th},	Second-order derivative (Jacobian matrix)
x ,	Random variable
x_t ,	Stochastic random process
(x_t) ,	Sequence of random processes, $i=1,2,\ldots n$
$\mathcal{Y}_t$,	σ-algebra generated by the observation processes $(y_s;\ s \leqslant t)$
Φ ,	Differentiable Gaussian measure, $N(\bar{x}, \Gamma_t)$
ν ,	Set of nonanticipative admissible controls
$\{W\}$,	Wiener-process sequence
(m.s.e.) ,	Mean square error
RMS ,	Root mean square
(P.a.s.) ,	Almost surely with probability one

AN OPTIMAL CONTROL APPROXIMATION
FOR A CERTAIN CLASS OF NONLINEAR FILTERING PROBLEMS

1. INTRODUCTION

The behavior of natural phenomena in general does not follow strict linear deterministic laws. Modern technology, with its great refinement of instrumentation, has made it abundantly evident that the formulations of natural laws in which troublesome nonlinear terms are suppressed or neglected to achieve workable linearizations, often lead to faulty results or to sacrifices of precision, which are no longer tolerable. Consequently, the simple mathematical deterministic model must be modified. This can be achieved by using a stochastic nonlinear model representation where the dynamic behavior of a physical system is formulated in terms of the evolution of the state (x_t) of the system, under the influence of random disturbance (ξ_t), as a solution of a set of stochastic differential equations:

$$\dot{x} = f(t,x_t, \xi_t), t\varepsilon T \ , \qquad (1\text{-}1)$$

where f (.) is a nonlinear, real n-vector function;

x_t - a stochastic process n-vector which usually cannot be measured directly,

t - the time variable, $t\varepsilon[0, T]$,

ξ_t - a vector stochastic process to model the random disturbance.

To avoid confusion, the following notation will be adapted in the sequel. By writing x_t, one interprets x as a stochastic process, while $x(t)$ will mean that x is a deterministic function of time. Moreover, $(x(t))$, (x_t) will mean deterministic and stochastic sequences respectively.

An important special class of (1-1) is the stochastic differential equation with the random disturbance modeled as an additive white Gaussian noise β_t, which has well-known characteristics. Hence, a white-noise model is used here with

$$\frac{dx}{dt} = f(x_t, t) + G(x_t, t)\beta_t \ . \tag{1-2}$$

It is important to note here that this model may lead to mathematical difficulties in the case of classical integrals. However, we may replace the white-noise model, (1-2), by an equivalent "Wiener-process" model as follows:

$$dx_t = f(x_t,t)dt + G(x_t,t)dw \ , \tag{1-3}$$

which is interpreted in the Ito sense [L1]. More detailed discussion of the above equivalent formulations is given in Appendix A.

In most practical problems, the states of the system, which model its dynamical behavior, connot be observed directly; only noisy versions can be measured. This noise is due to the unavoidable measurement error, and in most cases is modeled for convenience either as an additive Gaussian process or as an additive Wiener process. Consequently, a filtering algorithm may be employed to extract and estimate reliably the states of the system from the measured values. Stated another way, a given system has its dynamics modeled by the following version of (1-1):

$$\frac{dx}{dt} = f(t,\ x_t,\ y_t,\ \xi_t)\ , \tag{1-4}$$

where f(.), t, x_t, ξ_t are as defined before. Here y_t, the observable states, (can be measured directly) evolve according to the following nonlinear differential equations:

$$\frac{dy}{dt} = H(t,\ x_t,\ y_t,\ v_t)\ . \tag{1-5}$$

Again, H(.) is a nonlinear, real n-vector function, and v_t is a vector stochastic process modeling the observation noises, which may be caused mainly by measurement noises and environmental effects.

Assume that (1-4), (1-5) can be solved for each realization of ξ,v. Then, realization of x,y which are also of a stochastic nature, are defined. Furthermore, to completely define the filtering problem, a performance criterion must be stated, which defines in what sense the estimate should follow the state. In this research, the mean-square criterion is adapted because of both its mathematical convenience, and the distance measure it provides the estimate to follow "closely" the original state. Hence,

$$J(\bar{x}_t, x_t) = E\int_0^T (||x_t - \bar{x}_t||^2)dt \tag{1-6}$$

Here, $E(.)$, $||.||$, $\bar{x}_t$ are the expectation, the Euclidean norm, and the "best" estimate of x_t respectively. This particular criterion was chosen to ensure a "global" filtering criterion. However, other criteria might be used depending on the application of interest.

The filter equation can be modeled as:

$$\bar{x}_t = K\left(y(s);\ s\varepsilon[0,t]\right)\ , \tag{1-7}$$

where $\bar{x}_t$, the estimate, is defined once the structure of the K operator is obtained. Thus, if we assume that $E(||x_t\ ||^2)<\infty$, $t\varepsilon[0,T]$, then it is well known [D2] that $\bar{x}_t$, the best estimate, is generated by the following formula:

$$\bar{x}_t = E\left(x_t/y_s;\ s\varepsilon[0,t]\right)\ , \tag{1-8}$$

where $E(/)$ denotes the conditional expectation. Moreover, with additional assumptions about the structure of (1-4), and (1-5), a recursive version of (1-8), (closed form) can be found (i.e. the estimated value of $\bar{x}_{t+\Delta t}$ at time $(t+\Delta t)$ can be generated by a recursive formula if given the value of the estimate $\bar{x}_t$ at time t and the observations $(y_s;\ s\varepsilon(t,\ t+\Delta)))$. This is immediately recognized within the class of linear systems with linear observations and additive Gaussian noises, as the optimal (in m.s.e), state estimator and consists of a finite dimensional linear filter. The latter is due to the Gaussian assumption in the system which permits the conditional probability density function to be completely characterized by only the conditional mean and the conditional covariance [D2]. However, for nonlinear systems this fortunate situation does not generally exist. Obviously, nonlinear filtering techniques are more general and greatly expand the range of practical problems which can be handled. But, its optimal estimator generally consists of an infinite-dimensional system of moments, or equivalently a partial differential equation that has an infinite number of dimensions. Consequently, approximation and ad-hoc techniques usually are employed to construct practical filters for nonlinear systems. The classical methods described in recent literature for realizable nonlinear filters can be roughly classified into two categories. They are either probabilistic or statistical approaches. In the statistical approaches, the basic idea is to linearize the nonlinear equation; then the Kalman-Bucy method is applied. The linearization, which is a first order power-series expansion of the nonlinear terms, is generally performed either with respect to a given "reference trajectory" or with respect to an estimator that is obtained by filtering, such as the case in the Extended Kalman Filer EKF. But, these approaches have some drawbacks. For instance, in the first case the true trajectory should be close to the assumed one which is sometimes hard to fulfill. In the second case, since the filter parameters are functions of the state estimate, an error in the estimate impacts filter gain and can result in filter bias, inconsistency, and even divergence. In the probability approach, which was started by Stratonavich [S1] and developed by Kushner [K4], Bucy [B1], Jazwinski [J1],--to mention a few, the general procedures are as follows: First, the equation of evolution for the conditional probability density functions are determined. Then, equations of the conditional moments are developed. Finally, various heuristic assumptions and arbitrary truncation schemes are applied to the evolving infinite dimensional filter equations to generate finite versions.

The basic common assumption that is used in the filtering algorithms mentioned above, is the requirement of the "differentiability" or "smoothness" of the nonlinear terms. This results in replacing the global properties of the filter by local properties, and "derivatives" which are further aggravated by noise. This assumption restricts the range of direct application (unless they are heuristically modified) of the above algorithms.

It is well known that linear filtering is of paramount importance to sonar (active and passive) and radar applications. Estimation in sonar is often associated with the localization of a target which has already been detected. Localization is essentially a parameter estimation problem where the parameters of interest typically are target range, Doppler (radial velocity) and azimuth angle. The sonar "measures" one or more of these quantities as a function of time using the observed sensor data to obtain a history of the target track for surveillance or fire control. In general, the measured quantities are nonlinear functions of the localization parameters so that nonlinear estimation techniques must be used to establish the target track. Various manual - deterministic and automated - sequential methods such as the EKF are currently in use. There are obvious shortcomings of the EKF for passive tracking, such as the "ill conditioning" of the error covariance matrix due to false observability which causes the filter divergence. Detailed studies of the EKF shortcomings are given in [A2], and [C1]. Consequently, the development of a "global" or "local time" nonlinear filter that does not suffer from the above shortcomings would be quite attractive, specifically to sonar applications. The proposed nonlinear filters developed herein do have the above features. Hence, applicability to sonar applications must be fully explored.

This dissertation introduces new finite-dimensional filtering structures for a certain class of nonlinear systems which offers:

1. Better filtering accuracy than the traditional techniques.
2. Does not require the "smoothness" assumption of the existing techniques.
3. A global filtering criterion.

It is the general goal of this work to devise a means of coping with many of the drawbacks of the exisiting techniques. The object of this thesis is the design of a new nonlinear filtering approximation. The filter structure is recursive, easily implementable, efficient, and finite dimensional. The concept of both conditionally Gaussian processes and an optimal control approach to filtering are used in the filter derivation. The method of solution is based on a result of Liptser and Schiryayev [L2] which was rigorously extended to the vector case by Kolodziej [K1]. This approach combines the advantages of a sound theoretical basis, generation of an approximation model for the original nonlinear system, and generation of a finite-dimensional nonlinear filter which has certain improved characteristics over most classical methods, such as the popular EKF.

The suggested technique, which could be called "an approximation in the parameter space," consists of three basic concepts: approximate feedback control modeling, conditionally Gaussian filtering, and control law computation. The use of the conditionally-Gaussian concept, which was formally introduced by Lipster and Shiryayev [L2] allows a closed system of equations for calculating recursively $\bar{x}_t = E(x_t/y_t)$ and $\Gamma_t = \text{Cov}\ (x_t/y_t)$ which completely characterize the conditional Gaussian-distribution function $P(x_t \leq a/y_t)$, $t \geq 0$. That provides a certain class of stochastic nonlinear systems with the same tools as the Gaussian assumption provides for linear systems.

The main advantage of this approach over the traditional techniques are: First, the basic and most common assumption of the "differentiability" or "smoothness" of the nonlinear terms is not required here. This is important for filter stability and susceptability to noise aggravation; second, the model itself is approximated rather than approximating the evolving filter equations, as is the case in most of the existing filter methods. Consequently, one has more flexibility in adjusting the approximating parameters and end up with a good (in m.s.e. sense) approximation model, as well as an approximated finite-dimensional nonlinear filter. That is not the case when the filter equations are approximated directly, due to the strict requirements of the filter theory. Finally, in most linearization techniques the available observations are used only through the innovation process in the filtering algorithm so valuable information may be wasted. Herein, more complete use of the available observations are undertaken through the concept of conditioning the process on the given information as well as channeling it back through the innovations process.

The dissertation organization is as follows: chapter one is mainly tutorial in the sense that it reviews briefly the significant techniques that are used in nonlinear filtering theory. It also includes the definition of the goals, the scope of this research, and an outline of the organization of the dissertation. In chapter two, a formal definition of the nonlinear filtering problem is given, and the structure of an optimal finite-dimensional nonlinear filtering approximation is developed. In chapter three, a major feature, a "decompositon scheme," is incorporated in design strategy of filtering a certain class of large-scale nonlinear systems. Chapter four presents application examples to illustrate the proposed filtering algorithm via digital computer simulations. A comparison with traditional techniques such as the EKF and the Modified Truncation Second-Order Filter MSOF (when it is applicable) are also given. This comparison is based on performance, accuracy and cost of computation and storage requirements. Finally, Chapter five presents concluding remarks, and comments on future areas of research.

Appendix A presents a brief discussion of the equivalence between the Wiener-process formulation and the white-noise formulation in modeling stochastic differential equations. Appendix B is intended to be an easy reference to two Lemmas used in this text.

Chapter 2

2. AN APPROXIMATE OPTIMAL FINITE-DIMENSIONAL NONLINEAR FILTER

In this chapter the nonlinear filtering problem is defined, and the derivation of a bilinear feedback-control-model approximation with its corresponding finite-dimensional filter is presented. In general, the suggested technique uses extensively three basic concepts, bilinear feedback modeling, control-law computation, and conditionally-Gaussian filtering. Hence, these three significant concepts are formally introduced. The new filter, in general, consists of two major steps. In the first step, the given system is approximated by a bilinear, feedback-control model. In the second step, the state estimator (m.s.e.) is computed using the conditionally-Gaussian-filter format. Thereafter, it will be referred to as the two-step nonlinear filter (TNF).

2-1 A Nonlinear Filtering Problem

Consider as given some complete probability space (Ω, F, P) with a nondecreasing, family of sub-σ-algebra $\zeta_t \leqslant F$, $t \varepsilon [0,T]$. Let (w_{ti}), $i = 1,2,\ldots,n$ be mutually independent Wiener processes comprising a vector of dimension n. Also, given a nonlinear model with dynamics described by a family of stochastic differential equations (in the Ito sense) of the form

$$dz_t = f(t,z)\,dt + \sigma(t,z)dw_t \quad . \qquad (2\text{-}1)$$

Here, $t \varepsilon (0,T)$, $z \varepsilon C$ the space of continuous functions $< R^n$, and

$$f : [0,T] \times R^n \rightarrow R^n \quad \text{nonlinear, real n-vector functions,}$$

$$\sigma : [0,T] \times R^n \rightarrow R^n \times R^n \quad \text{matrix.}$$

The nonlinear functions involved in equation (2-1) are assumed to have appropriate properties to guarantee the existence, uniqueness, and continuity of sample solution with probability one (L1). These sample solution properties are crucial for modeling a physical practical system by (2-1) and essential for simulating the corresponding model by digital computer.

Suppose now that $z \varepsilon R^n$ is written in terms of two components $z=(x,y)$, where $y \varepsilon R^m$, $x \varepsilon R^{n-m}$. Correspondingly, f, σ can be written as

$$f = (F,H), \ \sigma = \begin{pmatrix} G \\ R \end{pmatrix},$$

where F,H are (n-m), m dimensional vectors respectively. G,R are of dimensions $(n-m) \times n$, $m \times n$ respectively. The vector y represents the noisy partial observations that are made of the whole process. Thus, the observation σ-field is the sub σ-field of z_t defined by $Y_t = \sigma\text{-algebra } (y_s;\ 0 \leq s \leq t)$. Notice that if m = n the system in (2-1) is completely observable. Hence, the system in (2-1) can be decomposed as follows:

$$dx_t = F\ (z,t)\ dt + G\ (z,t)\ dw_t\ , \qquad (2\text{-}2\text{-}a)$$

$$(2\text{-}2)$$

$$dy_t = H\ (z,t)\ dt + R\ (z,t)\ dw_t. \qquad (2\text{-}2\text{-}b)$$

Here again (2-2-a), (2-2-b) are interpreted in the Ito sense (L1).

The system in (2-2) is nonlinear, then the optimal estimate for a minimal-variance criterion, (i.e. mean-square-error sense), is known to be the conditional expectation of the state of the system given the observation, $\{y_s;\ 0 \leqslant s \leqslant t\}$, $t \varepsilon [0,T]$.

$$\bar{x}_t = E\ (x_t/y_s;\ s\varepsilon\ (0,t))\ , \qquad (2\text{-}3)$$

where $E(/)$ denotes the conditional expectation. In principle a sequential version of (2-3) can be found, (i.e. the estimate $\bar{x}_{t+\Delta t}$ at time $(t + \Delta t)$ can be generated recursively if given the value of the estimate $\bar{x}_t$ at time t and the observations $(y_s\ ;\ s\varepsilon[t,t+\Delta t])$. But, in general, the recursive formulae consist of an infinite-dimensional system of moment equations which are needed to completely characterize the conditional probability density $p(x_t, t/y_s\ ;\ s\varepsilon[0,t])$. Thus, an approximation must be made for practical implementation. A new approximation method will be developed in the following sections.

2-2 A Bilinear Feedback Approximation Model

A bilinear control model which approximates the nonlinear system in (2-2) is presented here. This technique is called "an approximation in the model parameter space." Its parameters are functions of the feedback control law u_t which is itself a functional of the observation processes y_t.

In general, the model can be approximated as follows:

$$\begin{aligned} dx_t &\approx \bar{F}(x_t,u_t,t)\, dt + \bar{G}\,(x_t,\, u_t,t)\, dw_t, \\ dy_t &\approx \bar{H}\,(x_t,u_t,t)dt + \bar{R}\,(x_t,u_t,t)dw_t\ , \end{aligned} \qquad (2\text{-}4)$$

where $(\bar{F}(.),\ \bar{G}(.),\ \bar{H}(.),\ \bar{R}(.))$ are functionals of u_t. The control u_t is measurable with respect to Y_t=σ-algebra $(y_s;\ 0\leqslant s\leqslant t)$, the stochastic process defined on $t\varepsilon[0,T]$, and it is chosen to minimize the following "mean-square global" criterion:

$$Q(u) = E(\int_0^T ||F_t - \bar{F}_t||^2\ dt)\ , \qquad (2\text{-}5)$$

where F_t denotes any of the functions F, H, R, or G in (2-2), and $\bar{F}_t$ is its corresponding approximation in (2-4). Here, the norm $||.||$ is a Euclidean norm, and the arguments $(x_t,\ y_t,t)$, (x_t,u_t,t)

are omitted for brevity. It should be noted that the choice of this performance index (2-5) ensures a form of global filtering even though, the search for the approximation $\bar{F}_t$ falls into a class of stochastic-control problems and depends strongly on the type of nonlinearities in the system. However, in [K3] it was emphasized that using a "local" filtering criterion will eliminate the dependence of the approximating parameters on the particular form of nonlinearities in the system. Thus, it is a trade off between filtering properties and the complexity of the search for the approximation parameters.

An important special class of (2-4) is the following bilinear form:

$$dx_t \simeq [A(u_t,t)x_t + B(u_t,t)]dt + \bar{G}(u_t,t)dw_t,$$
$$dy_t \simeq [C(u_t,t)x_t + D(u_t,t)]dt + \bar{R}(u_t,t)dw_t, \tag{2-6}$$

where $(A,B,C,D,\bar{G},\bar{R})$ are of appropriate dimensions, and are linear functionals of u_t. The term bilinear refers to the fact that the system is linear in the control and state, but not jointly linear [M2]. Now assume $\bar{u}_t$ which minimizes (2-5) is available; then (2-6) is a "close" approximation model to the original system in (2-2), and the minimization criterion is a measure of the quality of the approximation.

Remarks

(i) Although the bilinear structure in (2-6) requires the identification of more parameters than the ones in the original system in (2-2), it is mathematically more convenient for the derivation of the filter.

(ii) A unique, explicit, general form for $(A, B, C, D, \bar{G}, \text{and } \bar{R})$ cannot be given, because they depend strongly on the type of nonlinearities in the original system.

(iii) The model may include a broad class of non-Gaussian noise source, since non-Gaussian or even nonstationary processes may be modeled by Wiener noise passed through an appropriate nonlinear filter of this class.

(iv) The use of the feedback control in the approximation will couple the filter equations (i.e. the covariance equations will be functions of the estimates). But unlike the other linearization techniques this may enhance the stability of the filter due to the feedback structure.

2-3 Conditionally Gaussian Processes

An important concept which is used extensively in this research is the conditional Gaussian concept. Lipster and Shiryayev [L2] formally define it as follows:

Theorem 2-1:

Let (with probability one p.a.s) the conditional distribution $P(x_0 \leqslant a_0/y_0)$, $a_0 \varepsilon R$ be Gaussian, $N(m_0, \Gamma_0)$ with $0 \leqslant \Gamma_0 < \infty$. Then the random process (x_t, y_t), $0 \leqslant t \leqslant T$, satisfying a diffusion type of equations as in (2-6) where the parameters satisfying conditions (11-4)-(11-11) of [L2], is conditionally Gaussion such as: for any $t>0$, $0 \leqslant t_0 < t_1 \ldots . t_n \leqslant t$, the conditional distributions

$$F_{y_0}(x_0, \ldots, x_n) = P(x_{t_0} \leqslant a_0, \ldots, x_{t_n} \leqslant a_n / \zeta_t)$$

are (p.a.s) Gaussian.

The proof of this theorem is very long and is given in (L2). This result is very important since it allows a closed system of equations for generating recursively $\bar{x}_t = E(x_t/y_t)$, and $\Gamma_t = \text{Cov}\,(x_t/y_t)$. (This is obtained by replacing the complex computation of the conditional expectation in (2-3) by a simple integration.) And these two parameters completely characterize the conditional distribution $P(x_t \leqslant a/y_t)$, $t \geqslant 0$. So, it provides a class of stochastic nonlinear systems with the same tools that the Gaussian assumption provides for linear systems.

Subsequently, for this class of nonlinear systems, the concept led to the development of a finite conditionally-Gaussian filter by Lipster and Shiryayev [L2]. It also, offers more flexibility in control applications than do linearization techniques. This advantage has been demonstrated by Kolodziej [K1] and Mohler and Kolodziej [M1].

It should be emphasized here that in application the necessary assumption of x_0 given y_0 to be conditionally Gaussian can be satisfied under realistic operating conditions. This may result from a physical consideration or from a direct approximation of the distribution of x_0 given y_0 by a Gaussian process. In the first case, for example, the error of the estimate of x_0 given y_0 might be caused by many random phenomena which in turn might be approximated by a Gaussian process. Nevertheless, this does not necessarily mean that either x_0 or y_0 has to be Gaussian.

2-4 A Conditionally Gaussian Filter

This finite-dimensional filter is derived by Lipster and Shiryayev [L2] and is rigorously extended to the multi-dimensional case by Kolodziej [K1] who also, relaxed some of the required conditions suggested earlier in [L2].

To summarize their results, consider the system in (2-6) which is partially observable. At any time t it is desired to estimate the unobservable state x_t using realizations of the observation state y_t. Let (x_0,y_0) be the initial states for (2-6) which are assumed to be independent of the Wiener processes $w_i, i=1,2$. The parameters $(A(.),\ B(.),\ C(.),\ D(.),\ \bar{G}(.),\ \bar{R}(.))$ are of the appropriate dimensions and their elements are assumed to be measurable non-anticipative functionals on $[0,T]xC_T{}^n$. Also, assume that u_t is a measurable functional of Y_t where Y_t is the σ-algebra generated by $[y_s;\ 0 \leqslant s \leqslant t]$. Then, sufficient conditions for derivation of a recursive optimal, mean-square estimate of x_t given $(y_s;\ 0 \leqslant s \leqslant t)$ are given below. For all $u \in C_T{}^n$,

$$\int_0^T ||\ A(t,u)||^2\ dt \leqslant \infty, \text{ and} \tag{2-7}$$

$$\int_0^T [||B(t)||^4 + ||C(t,u)||^4 + ||D(t,u)||^2 + \bar{G}(t,u)||^4] dt \leqslant \infty\ . \tag{2-8}$$

For $u,n \in C_t{}^n$, $t \in [0,T]$ define

$$R^2(t,u) = \bar{R}(t,u)\ \bar{R}^*(t,u)\ ,$$

$$\text{then } ||R^{-2}(t,u)|| \leqslant c < \infty\ . \tag{2-9}$$

Also, assume that

$$||R(t,u) - R(t,n)||^2 \leqslant c\ (\int_0^T ||u(s) - n(s)||^2 dk(s)$$

$$+ ||u(t) - n(t)||^2)\ ,$$

and (2-10)

$$||R(t,u)||^2 \leqslant c \left(\int_0^T (1 + ||u(s)||^2 \, dK(s) + 1 + ||u(t)||^2 \right),$$

where $K(s)$ is a nondecreasing right-continuous function $0 \leqslant K(s) \leqslant 1$; c is a positive constant. The following comments are in order:

(i) conditions (2-7) - (2-10) ensure the existence and uniqueness of a uniform parabolic solution to the system in (2-6) which is important for real system modeling and simulation by digital computer;

(ii) conditions (2-7), (2-8) are assumed to assert that (x_t, y_t) are square integrable. For example, $A(t,u_t) = x^2$ will violate these conditions, however $A(t,u_t) = \sqrt{X}$ will satisfy these conditions. These conditions also, imply that

$$\int_0^T ||B(t,u)||^4 dt < \infty \quad .$$

This is important since it will restrict the additive stochastic control to the square integrable class that satisfies $E(\int_0^T ||u_t||^4 \, dt) < \infty$;

(iii) condition (2-9) is made so that no degenerate stochastic measure will be associated with Y. Otherwise, no uniform parabolic solution exists for the system in (2-6) if a noise term is missing in the equations of system (2-6);

(iv) The conditions in (2-10) restrict the noise coefficients $\bar{R}$ in (2-6) to a class of smooth functions of u_t, (due to the first equation of (2-10) which is a generalized Lipschitz condition). And the second equation of (2-10) ensures that its rate of growth is limited to at most linear growth of u_t. This is important so that the solutions (x_t, y_t) do not "explode" in finite time.

From Theorem (2-1) of [K1] the following results are quoted:

Theorem 2-2

Let equations (2-6) have a weak solution, (see Appendix A for definition of strong and weak solutions), (x_t, y_t), $t \varepsilon [0,T]$ with the initial states (x_0, y_0) satisfying

$$E\,[\|x_0\|^4] < \infty\,,$$
$$P(\|\,y_0\| < \infty) = 1\,. \tag{2-11}$$

Let the conditional distribution $P(x_0 \leqslant \alpha_0/y_0)$ be (P.a.s) Gaussian with parameters $\bar{x}_0 = E(x_0/y_0)$, $\Gamma_0 = E[(x_0-\bar{x}_0)(x_0-\bar{x}_0)^*/y_0]$ and $tr(\Gamma_0) < \infty$ p.a.s. If conditions [(2-7) - (2-10)] are satisfied, then the processes (x_t, y_t), satisfying (2-6), $t\varepsilon[0,T]$ are conditionally Gaussian, i.e. for any $t\varepsilon[0,T]$ and any finite partition t_j, $j=0,1,\ldots k$ of $[0,T]$ such that $0 < t_0 < t_1, \ldots t_k \leqslant t$, the conditional distribution

$$P(x_{t0} \leqslant \alpha_0,\ x_{t1} \leqslant \alpha_1, \ldots x_{tk} \leqslant \alpha_t / y_t)\ ,\ \alpha \varepsilon R^n$$

is p.a.s. Gaussian, Y_t is σ-algebra generated by $\{y_s;\ 0 \leqslant s \leqslant t\}$. Further, $\bar{x}_t$, Γ_t, i.e. $\bar{x}_t = (x_t/y_t)$ and $\Gamma_t = E[(x_t - \bar{x}_t)(x_t - \bar{x}_t)^* / y_t]$, are unique continuous solutions to

$$d\bar{x}_t = (A\bar{x}_t + B)dt + Kdv,$$

$$K = (\Gamma_t C^*)\ (RR^*)^{-0.5},$$

$$dv = (RR^*)^{-0.5}(dy_t - (C\bar{x}_t + D)dt), \tag{2-12}$$

$$d\Gamma_t = (A\Gamma_t + \Gamma_t A^* + GG^* - KK^*)dt,$$

with $\bar{x}_0 = x(0)$, $\Gamma_0 = \Gamma(0)$ as initial conditions.

The proof of the theorem is parallel to the proof given in [K1] and [L1]. The interested reader is referred to the given references for details. The equations in (2-12) are recursive formulae (filter) for an optimal mean-square estimate of x_t. The arguments (t,u) are omitted again for brevity. The finite dimensionality of this filter is due to the conditional-Gaussian assumption of the processes (x_t, y_t). A schematic representation of the filter is given in Figure 2-1 to realize a functional understanding of its design.

2-5 A Stochastic Nonlinear Control Problem

Following Lipster and Shiryayev [L2] the partially observable system (2-6) is transformed into a completely observable system (2-12). Consequently the minimization criterion (2-5) also must be transformed in order to solve for the control law analytically. Let the $u_t = \psi(t,y)$ where $\psi:[0,T]xC\to R^k$ such that ψ is nonanticipative, (i.e. $\Psi(t,y) = \Psi(t,y^*)$ if $y(s) = y^*(s)$ for $s \leq t$). Now, if ψ is nonlinear, (x_t,y_t) which satisfy (2-6) will no longer be a normal process, but the conditional distribution of x_t given $\{y_s;\ s \leq t\}$ is still normal with mean and covariance given by (2-12).

Let $\Phi\ (\xi,t,\bar{x},\Gamma_t)$ denote a differentiable Gaussian measure with mean $\bar{x}$, and variance Γ_t.

Define

$$\bar{L}\ (t,\ \bar{x}_t,\ \Gamma_t,\ u_t) \equiv \int_R L\ (t,\ \xi,\ u)\ d\ \Phi\ (\xi,\ t,\ \bar{x},\ \Gamma_t)\ , \tag{2-13}$$

where L denotes $||F - \bar{F}\ ||^2$, and F, $\bar{F}$ are as defined before.

If Γ_t is nonsingular, then

$$d\ \Phi\ (\xi,\ t,\ \bar{x},\ \Gamma_t) = \frac{1}{(2\pi)^{n/2}|\Gamma_t|}\ \text{Exp} - 0.5\ [(\xi-\bar{x})^*\Gamma^{-1}(\xi-\bar{x})]\ d\xi\ , \tag{2-14}$$

where $|\Gamma_t|$ denotes the determinant of Γ_t, and Γ^{-1} is the inverse matrix of the covariance.

Now, using the properties of expectation, the Bayes formula, and the definitions in (2-13) and (2-14), the cost function in (2-5) can be written as

$$Q(u) \simeq E\left(\int_0^T E(||F - \bar{F}||^2/y_s;\ s \leq t)dt\right),$$
$$\simeq E\int_0^T \bar{L}(t, \bar{x}_t, \Gamma_t, u_t)\ dt. \tag{2-15}$$

The interchange of integration and conditional expectation in (2-15) is permitted by a version of Fubini's theorem [R1], and it is justified if we are dealing with integrals of Gaussian random processes.

Thus, an equivalent completely observable system (2-12), (2-15) emerges where the new states of the system $(\bar{x}, \Gamma_t)$ are generated by (2-12). But the parameters of this equivalent system are functionals of the control law. Consequently, we have to solve a control-law problem first. Thus, the filtering problem is actually replaced by a stochastic-control problem, which results in a stochastic Bellman equation. Apparently, this could lead to a more difficult problem to solve than the original filter problem for (2-2). But as has been shown in [K2] an approximation to the control-law can be found whithout solving the Bellman equation. In the mean time, the control-law problem is formulated as follows. Let $z = \begin{pmatrix} \bar{x} \\ \Gamma_t \end{pmatrix}$, a vector of appropriate dimensions. Then $dz = \begin{pmatrix} d\bar{x} \\ d\Gamma_t \end{pmatrix}$, and hence (2-12) can be written as

$$dz = f(z,u,t)dt + \sigma(t,u)dv \ , \tag{2-16}$$

where

$$f(z,\ u,\ t) = \begin{pmatrix} A(t,u)\bar{x} + B(t,u) \\ A\Gamma_t + \Gamma_t A^* + \overline{GG}^* - KK \end{pmatrix},$$

$$\sigma(t,u) = \begin{pmatrix} k(t,u) & 0 \\ 0 & 0 \end{pmatrix}.$$

Here the matrices are of appropriate dimensions. But equation (2-16) is a "degenerate" equation because the noise term is missing in the covariance equation. Thus, as pointed out in the remark about condition (2-9), no uniform parabolic solution exist for the system because $\sigma(t,u)\ \sigma^*(t,u)$ is singular which implies the existence of singular probability measure for $\bar{x}$, Γ_t. Hence, the approximation of a degenerate system of stochastic equation proposed by Fleming [F1] which satisfies condition (2-9) is adapted here, where small white-noise terms may be added to the covariance equation. Thus (2-16) becomes

$$dz_e = f(z,\ u,\ t)\ dt + \bar{\sigma}dw, \tag{2-17}$$

where $\overline{\sigma} = \begin{pmatrix} K(t,u) & 0 \\ 0 & (2e)^{0.5} \end{pmatrix}$, $dw = \begin{pmatrix} dv \\ dw \end{pmatrix}$, f(.) is as defined in (2-16).

For the stochastic control problem (2-15), (2-17), it has been shown by Fleming [F1], Davis [D1], and Ahmed and Tio [A1] that there exists an optimal control for the system described by (2-15), (2-17) if the following conditions are satisfied: Let $Q = (T_0 \times T)xR^n$ where $T_0 \leqslant T$;

(i) $\overline{\sigma}$, f are Lipschitz continuous in z, $(z = (\bar{x} , \Gamma_t)$;

(ii) $\overline{\sigma}$ is nonanticipative with respect to w_i, i = 1,2;

(iii) f, L are bounded measurable on $\bar{Q}$ (the closure of Q) for each $u \epsilon \upsilon$, where υ is the set of admissable controls which is continuous for every $(t,z) \varepsilon \bar{Q}$;

(iv) f, L are convex for each $(t,z) \varepsilon Q$;

(v) $a(t,z) = tr(\sigma\sigma^*)$ is nonsingular.

Using dynamic programming it is seen that there exists a value function $V(t, \bar{x}, \Gamma)$ which is differentiable at least once in t and twice in z, and which satisfies:

$$V(t,z) = \inf_{\upsilon(0,T)} \left(E_{t,\zeta} \int_t^T L(t, z, u)dt\right) . \tag{2-18}$$

The corresponding stochastic Bellman equation is

$$V_t + \min_{u} [0.5 \, tr(\sigma^* V_{zz} \sigma) + V_z^* f(z, u, t) + \bar{L} (z, u, t)] = 0,$$

and $\quad V(z, u, T) = 0$, (2-19)

where $\quad V_t = \frac{\partial V}{\partial t}$, $V_z = \frac{\partial V}{\partial z}$, $V_{zz} = \frac{\partial V^2}{\partial z_i \partial z_j}$.

Consequently, a function V(.) can be found satisfying equation (2-19), i.e., with corresponding control function $u^0(t,z)$, such that

$$\frac{\partial V(t,z)}{\partial z} f(t, z, u^o) + \bar{L} (t, z, u^o) = \tag{2-20}$$

$$\min_{u \varepsilon \nu} [\frac{\partial V}{\partial z} (t,z) \, f \, (t, z, u) + \bar{L} (t, z, u)],$$

then u^o is optimal. Futhermore, if conditions (i)-(v) are satisfied, then (2-17) has a smooth solution.

Intuitively, the optimal control is a function of the value function (the solution of the stochastic Bellman equation) but the computation (requires solving a nonlinear Cauchy-type problem) of the exact value of the optimal control is in general tremendously complex. However, suggested abstract methods to solve a Cauchy-type problem as in (2-17) will be discussed briefly.

The nonlinear Bellman partial differential equation (2-19) can be stated in general as a Cauchy problem of the following form:

$$C_t + H(t,\ x,\ C_t,\ C_{xx}) = 0, \tag{2-21}$$

$$C(0,x) = C_0,$$

where C_t denotes the partial derivative with respect to time, and C_x, C_{xx} the gradient and the Jacobian respectively.

Roughly speaking, there are three principle approaches for solving the problem in (2-21):

(1) Separation of variables (also called the Fourier method, or solution by eigen-function expansion).

(2) Green function (also called fundamental singularities, or solution by integral equation).

(3) Variational formulation (also called the calculus of variations).

Accordingly, some of the important methods that have been discussed in the recent literature are:

(1) The parametric method, which is developed by Friedman [F2], [F3]. Here a fundamental solution is first constructed, then used to solve the Cauchy problem.

(2) Hilbert-space method [G1]. The idea of separation of variables in the context of Hilbert space, and numerical methods are used. This is a very elegant method but unfortunately it is extremely limited in its scope of application.

(3) Function theoretical method [G2], where the theory of integral operators is used. But there is no unique method to determine the integral operator, and hence a general solution.

(4) Characteristic method, [B2], where the idea of reducing the P.D.E. problem to an O.D.E. problem, through the use of the corresponding characteristic function. Then a set of Hamilton's equations is solved by simple integration. This method guarantees a solution only where $(t;x) \rightarrow X$ is invertible, that is, as long as the characteristics do not intersect. But any nonlinearities in the Hamiltonian lead to crossing of characteristics. Hence, the application to nonlinear systems is limited.

(5) Transformation methods [A3], where the transformation may be applied to either the dependent variable, the independent variables, the equation itself, or any combination of these. However, the initial-condition transformation may be a real problem in this technique. For example, if the equation in two independent random variables $tc_t + xc_x = 2txc$, $c(0,x) = a$, is transformed by $v = tx$ into the ODE, the single variable v can take on only a single value on any manifold tx = constant.

The choice of method of solution to (2-21) from the above suggested techniques, depends largely on the following factors:

(i) degree of nonlinearities in the system, (i.e. some of the methods may not be applicable);

(ii) degree of difficulties and complexity of the method.

However, the transformation method seems to be the most promising one due to its simplicity and the existence of various transformation techniques (e.g., the cononical transformation and the similarity methods). For example, if the system in (2-19) is first order, then the following method transforms the system to a quasilinear system.

$$C_t = -H\ (t,\ x,\ C_x), \qquad (2\text{-}22)$$

$$C(0,x) = f(x)\ .$$

Let P denote C_x and $q = P_o = C_t$. Then by differentiation with respect to t and equating mixed partial derivatives, the following quasilinear system is obtained:

$$\begin{aligned} &C_t = q, \\ &P_t = q_x, \qquad (2\text{-}23) \\ &q_t = -H\ (t,\ x,\ P) - H_p\ (t,\ x,\ P)\ q_x, \\ &C(0,x) = f(x), \\ &P(0,x) = f_x(x), \\ &q\ (p,x) = -H\ \left(0,x,f_x(x)\right) \end{aligned}$$

That is if C is a twice continuously differentiable solution of the Bellman equation, then (C, P, q) is a solution of (2-22), and the converse can be proven to be true also.

2-6 An Optimal Finite-Dimensional Filtering Algorithm

The general filtering algorithm consists of the following steps:

Step 1 - Approximate the given nonlinear system, (i.e. the system in (2-2), by a bilinear feedback model that has the form in (2-6), where the parameters are linear nonanticipative functions/ functionals of the control u_t, (which is measurable with respect to the σ-algebra $Y_t = \{y_s;\ s \leqslant t\}$.

Step 2 - Choose the performance index for the global filtering criterion as in (2-5), which is also a measure of the quality of the model approximation in (2-6). If $\bar{u}_t$ which minimizes (2-5) is obtained, then (2-6) is a "close" model approximation to (2-2).

Step 3 - If the parameters (A, B, C, D, $\bar{G}$, $\bar{R}$) satisfy the conditions in (2-7)-(2-10), then the corresponding finite-dimensional filter has the form as in (2-12), with its parameters as functions or functionals of $\bar{u}_t$.

Step 4 - Transform the performance index defined in Step 2 to an equivalent criterion as in (2-15) using (2-13), and (2-14).

Step 5 - The system in (2-15), and (2-17) forms a completely observable stochastic control problem, which should be solved by classical dynamic programming to find the optimal control $\bar{u}_t$.

Step 6 - Once the optimal control $\bar{u}_t$ is obtained from Step 5 and then substituted back into both (2-6) and (2-12) a model approximation as well as a finite-dimensional nonlinear-filter approximation will be generated.

Remarks:

(i) The assumption that the solution of the control problem in Step 5 is close to $\bar{u}$, which minimizes the performance criterion defined in Step 2, resembles the assumption in the EKF approach, and closes the approximation procedures.

(ii) In general, the stochastic-control problem in Step 6 results in the nonlinear Bellman partial-differential equation as in (2-19). Hence, it may be no simplification as compared to the original problem of finding a filter for (2-2). This is an indication of a possible problem one will encounter in using this approach if the exact value of the controls are required.

However, as was emphasized in [K2], the structure of the optimal control suggested by an approximate solution to the Bellman equation, which results in an approximate control that yields a better performance filter than the EKF.

The following example [K2] illustrates the previous filtering algorithm steps:

Let
$$dx_t = a\,|x|\,dt + dw_1,$$
$$dy_t = x_t dt + dw_2. \tag{2-24}$$

The bilinear feedback approximation model is

$$dx_t \approx \bar{u}_t \ x_t dt + dw_1, \tag{2-25}$$

$$dy_t \approx x_t dt + dw_2; \ t\varepsilon[0,T],$$

where $\bar{u}_t$ is chosen to minimize

$$Q(u) = \min_{\bar{u}_t} E(\int_0^t (a \ |x| - \bar{u} \ x_t)^2 dt \ . \tag{2-26}$$

From (2-12), the optimal (m.s.e)-filter equation has the following form:

$$d\bar{x}_t = \bar{u}_t \bar{x}_t dt + \Gamma_t dv_t, \tag{2-27}$$

$$d\Gamma_t = (2\bar{u} \ \Gamma_t + 1 - \Gamma_t^{\ 2}) \ dt,$$

where $dv_t = dy_t - \bar{x}_t dt$, is the innovation process, having relevant properties of the Wiener process. Now from (2-13), (2-14) the equivalent minimization criterion to (2-26) is

$$\bar{Q} \ (t, \ \bar{x}, \ t) = \min_{u} E \int_0^T (a \ |\xi| - \bar{u}\xi)^2 \ d\Phi(\xi), \tag{2-28}$$

where $d\Phi(\xi)$ is as defined in (2-14).

Using the dynamic-programming approach for the stochastic control system defined by (2-27), (2-28) one can show [D2] that the optimal $\bar{u}$ is given by

$$\bar{u} = a \left[\frac{2\bar{x}\Gamma}{\sqrt{2\pi\Gamma}} \exp\left(-\frac{\bar{x}^2}{2\Gamma}\right) - 2(\bar{x}^2 + \Gamma) \operatorname{erf}\left(-\frac{\bar{x}}{\sqrt{2\Gamma}}\right) \right. \tag{2-29}$$

$$\left. - \left(\Gamma\frac{\partial V}{\partial \Gamma} + 0.5\, \bar{x} \frac{\partial V}{\partial \bar{x}}\right)\right] (\bar{x}^2 + \Gamma)^{-1},$$

where $\operatorname{erf}(\alpha) = \int_0^{\alpha} \frac{1}{\sqrt{2\pi}} e^{-0.5\xi^2} d\xi$.

Here, V is the solution to the Bellman equation

$$\frac{\partial V}{\partial t} + 0.5\, \Gamma^2 \frac{\partial^2 V}{\partial \bar{x}^2} + \frac{\partial V}{\partial \Gamma}(1 - \Gamma^2) + (\bar{x}^2 + \Gamma)\left(a^2 - \frac{\bar{u}^2}{a^2}\right) = 0,$$

$$V(T, \bar{x}, \Gamma) = 0. \tag{2-30}$$

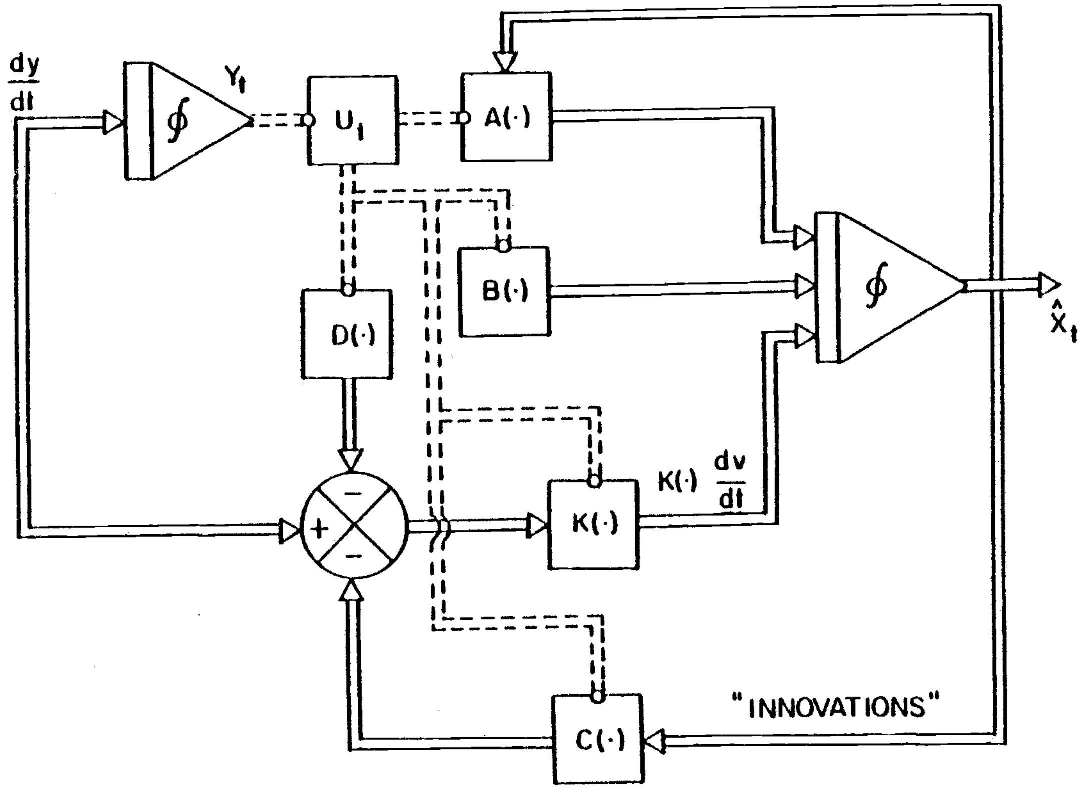
The above partial differential equation is very difficult to solve analytically, and it is a good indication of the possible problem one may face while attempting to solve the stochastic problem defined in Step 6 of the algorithm. However, an approximate solution to (2-30) of the form

$$\bar{V}(\bar{x}, \Gamma, t) = p(t) \exp\left(-\frac{\bar{x}_t^2}{2\Gamma_t}\right), \quad p(T) = 0, \tag{2-31}$$

when substituted into (2-29) yields the approximate control

$$\tilde{u} = a\left(\frac{2\bar{x}\sqrt{\Gamma}\,\exp(-\frac{\bar{x}^2}{2\Gamma})}{\sqrt{2\pi}\,(\bar{x}^2 + \Gamma)} - 2\,\mathrm{erf}\,(-\frac{\bar{x}}{\sqrt{2\Gamma}}\right) , \qquad (2\text{-}32)$$

which as shown in [K2] gives better performance than the EKF.



NOTE: The double lines are for vector representation. The symbol $\oint$ refers to stochastic integration. ⟹ Vector multiplication.

= = ⊃ The dashed lines implies functional representation.

Figure 2-1

Schematic Diagram for the Proposed Conditionally-Gaussian Filter

CHAPTER 3

3. AN APPROXIMATE NON-LINEAR - FILTER STRUCTURE

In this chapter a new filtering structure for approximating a certain class of large-scale nonlinear filtering problems is presented. A major feature, i.e., a decomposition scheme is incorporated into the filter design. The motives to develope this scheme are: one, to resolve the "curse" of dimensionality encountered if an optimal approximations is sought, due to the requirement of solving a control problem which is a function of $(\frac{n + n\ (n+1)}{2})$ variables. In general, this is a very difficult problem to solve. Two, to alleviate the difficulties with respect to control policy definition and calculation, which arise if the (TNF) is applied directly and the global properties are assumed, but an approximate control is sought. Finally, using this scheme will reduce the complexity of the algorithm, and results in significant computer saving in the digital simulation.

The strategy adapted in this decomposition scheme is based on the decomposition of the system into two inter-connected subsystems. The definition of the subsystem can be imposed using purely physical reasoning. The first subsystem is only linear and it will act as a first stage of the filtering process. The second subsystem includes all the nonlinearities in the system, which are then approximated by the proposed two-step nonlinear filter (TNF). This scheme has the following advantages:

(i) only the linear filtering problem is solved in the first level;

(ii) the control parameter which is needed in the TNF algorithm, will be easy to obtain as a function of the parameters of the first stage linear filter. So, no difficulties with respect to measure theory or control policy will be encountered if the global filtering criterion is retained;

(iii) in many cases, there is a substantial computational saving as compared to the global single system solution.

An interesting class of nonlinear system has the following form

$$
\begin{aligned}
dx_t &= F(x,y,t)dt + G(x,y,t)dw, \\
dy_t &= H(x,y,t)dt + \sigma(t)dv \ ,
\end{aligned}
\tag{3-1}
$$

where w,v are mutually-independent vector Wiener processes of appropriate dimensions; $\sigma(t)$ is a matrix of compatible order. The functions F(.), H(.), and G(.) can be partitioned as follows:

$$
F(x,y,t) = f_1(t)\,x + f_2(x,y,t) \ , \tag{3-2}
$$

$$
H(x,y,t) = \begin{pmatrix} h_1(t)x \\ h_2(x,y,y) \end{pmatrix} ,
$$

$$G(x,y,t) = g_1(t) + g_2(x,y,t) \quad .$$

Here, f_1, h_1, and g_1 are linear matrices of appropriate dimensions and f_2, h_2, and g_2 are nonlinear functions of their arguments, and of compatible orders.

The general outline of the decomposition scheme and filtering algorithm are: (1) the given n-dimensional system as in (3-1) is decomposed into two subsystems. Subsystem I consists of a linear system, (i.e. linear dynamic and observation equations). Subsystem II is a nonlinear system, which contains all the nonlinearities of the original system and is approximated by the proposed model space approximation.

(2) Apply a classical filtering technique, i.e. the Kalman-Bucy algorithm [D2] to the linear system in Subsystem I. This will be considered as the first stage of the filtering algorithm.

(3) Find an appropriate bilinear approximation model to the nonlinear system in Subsystem II.

(4) Finally, from Subsystem I and the bilinear approximation of Subsystem II, form a new system. This new system will be of the form that has a finite-dimensional, conditionally-Gaussian filter as in (2-12) if certain assumptions (2-7)-(2-10) (See Chapter 2 Section (2-3).) about the system parameters are satisfied.

The block diagram representation of the decomposition scheme and the filtering algorithm is given in Figure (3-1), while a detail schematic representation is given in Figure (3-2).

To summarize, the various algorithmic steps are:

Step 1. The given nonlinear system as in (3-1) can be decomposed into two subsystems.

Subsystem I

$$dx_{ft} = f_1(t)\ x\ dt + g_1(t) dw \quad ,$$

$$dy_{ft} = h_1(t)\ x\ dt + \sigma_1(t)\ dv_1 \ , \tag{3-3}$$

where f_1, h_1, and g_1 are as defined before, and w, v_1 are independent Wiener processes of appropriate dimensions. The subscript f denotes the first subsystem.

Subsystem II

$$dx_{st} = f_2(x,y,t)dt + g_2(x,y,t)dw \quad , \tag{3-4}$$

$$dy_{st} = h_2(x,y,t)dt + \sigma_2 dv_2 \ ,$$

where the subscript s denotes the second subsystem, $f_2(.)$, $h_2(.)$, and $g_2(.)$ are generally nonlinear functions of their arguments, and w, v_2 are independent Wiener processes of compatible orders.

Step 2 The Kalman-Bucy filter equations for the system in (3-3) are:

$$d\bar{x}_{ft} = f_1\bar{x}_f dt + ph_1^*(\sigma_1\sigma_1^*)^{-1} dn_t \,, \qquad (3\text{-}5)$$

$$dp_t = (f_1 p + pf_1^* + g_1 g_1^* - ph_1^*(\sigma_1\sigma_1^*)^{-1} h_1 p)\, dt \,,$$

$$dn_t = dy_{ft} - h_1\, \bar{x}_{ft}\, dt \,,$$

$$\bar{x}_f(0) = E(x_f(0)) \,, \; p(0) = \mathrm{cov}\,(x_f(0)) \,,$$

where $\bar{x}_f$ is the estimate, and $p(t)$ is the error covariance matrix. Here dn_t is the innovation process.

Step 3 The bilinear approximation model for the system in (3-4) has the following form:

$$dx_s \approx \bar{f}_2\,(t, u_t, x)\, dt + \bar{g}_2\,(t, u_t)\, dw \,,$$

$$dy_s \approx \bar{h}_2\,(t, u_t, x)\, dt + \sigma_2 dv_2 \,, \qquad (3\text{-}6)$$

where

$$\bar{f}_2\,(x, u_t, t) = \sum_{i=1}^{n} u_i(t)\, x_i(t) + u_{n+1}(t) \quad = A(t,u)\, x + B(t,u) \,,$$

$$\bar{h}_2(x,u_t,t) = \sum_{j=1}^{n} u_j(t)\, x_j + u_{n+1}(t) \qquad = C(t,u)\, x + D\,(t,u)\ ,$$

$$\bar{g}\,(t,u_t) = g_0\,(t,\ u_t)\ . \tag{3-7}$$

Here n, the dimension of the system, and the second equality, is used for mathematical convenience. The control $u_1(t)$, $1 = i = j = 1,2$-- $n+1$ are measurable with respect to σ-algebra $\{y_{fs}; s\varepsilon[0,t]\}$, and are chosen to minimize the following global filtering criterion:

$$Q(u) = \min_{u_1} E \int_0^T (k-\bar{k})^2 dt. \tag{3-8}$$

Using the property of expectation and Bayes formula, (3-8) becomes

$$Q(u) = \min_{u_1} E \left(\int_0^T E(k-\bar{k})^2/y_{fs};\ s \leqslant t)\ dt\ \right)$$

$$= \min_{u_1} E \left(\int_0^T E^t (k-\bar{k})^2 dt\right)\ , \tag{3-9}$$

$$= \min_{u_1} E \left(\int_0^T H(x,u)\ dt\ \right)\ ,$$

where E^t denotes the conditional expectation. Here again, k denotes any of the functions f_2, h_2, or g_2, while $\bar{k}$ denotes the corresponding approximation $\bar{f}_2$, $\bar{h}_2$, or $\bar{g}_2$ in (3-7), and $H(x,u) = E^t(k-\bar{k})^2$. The arguments (t,y,x), (t, u_t) are omitted for brevity.

The minimization of (3-9) with respect to u_l, $l=1,2,\ldots n+1$, can be performed since the expectation is conditioned on the σ-algebra $(y_{fs}$; $s\varepsilon[0,T]$, which is not a function of the parameters u_l, $l = 1,2\ldots n+1$. Moreover, the parameters u_l, $l=1,2\ldots n+1$ are functions of the states x_i, $i=1,2\ldots n$ but the states x_i, $i=1,2\ldots n$ are not function of the parameters u_l, $l=1,2\ldots n+1$. In fact, the problem can be treated as a special case of an open-loop control problem formed by dynamics (3-6) and cost (3-9), because as far as the newly formed problem is concerned the parameters u_l, $l=1,2\ldots n+1$ are functions of $t, t\varepsilon[0,t]$. According to the argument given in Chapter 2-Section (2-5) and [D2], u_l can be found by solving the following equation:

$$V_t + 0.5\mathrm{tr}\,(\sigma V_{xx}\sigma^*) + \min_{u_l}\,[V_x f(x) + H(x,u)] = 0,\ l = 1,2,\ldots n+1, \tag{3-10}$$

where f(.) denotes any of the functions f_2, h_2, or g_2, and V_t, V_x, are as defined in (2-19). But as pointed out earlier, f(x) are not functions of u_l $l=1,2,\ldots n+1$,

$$\text{thus } \frac{\partial f(x)}{\partial u_l} = 0 \ ,$$

and (3-10) can be written equivalently as

$$\min_{\bar{u}_l} [H(x,n)] = 0, \; l=1,2,\ldots n+1 \, . \tag{3-11}$$

Now, to obtain the parameters $\bar{u}_l$, $l=1,2,\ldots n+1$ which minimizes (3-9), a set of equations resulting from (3-11) must be solved simultaneously and as pointed out previously, they are functions of $(\bar{x}_f, p)$.

Step 4 The new equivalent system has the following form:

$$\begin{aligned} dx_t &\approx \left(A_1(t,\bar{u}_t)\, x_t + B_1(t,\bar{u}_t)\right) dt + G(t,\bar{u}_t)\, dw, \\ dy_t &\approx \left(C_1(t,\bar{u}_t)\, x_t + D_1(t,\bar{u}_t)\right) dt + \sigma dv, \end{aligned} \tag{3-12}$$

where

$$A_1(t,\bar{u}_t) = f_1(t) + A(t,\bar{u}_t), \; B_1(t,\bar{u}_t) = B(t,\bar{u}_t),$$

$$C_1(t,\bar{u}_t) = \begin{pmatrix} h_1(t) \\ C(t,\bar{u}_t) \end{pmatrix}, \; D_1(t,\bar{u}_t) = \begin{pmatrix} 0 \\ D(t,\bar{u}_t) \end{pmatrix},$$

$$G_1(t,\bar{u}_t) = g_1(t) + g_0(t, \bar{u}_t), \quad \sigma = \begin{pmatrix} \sigma_1(t) & 0 \\ 0 & \sigma_2(t) \end{pmatrix} .$$

Here again the matrices are of compatible orders.

Step 5 With certain assumptions about $(A_1, B_1, C_1, D_1, G_1, \sigma)$ and the distribution of the initial state x_0 given y_0 (see Chapter 2 Sections 2-3), the corresponding conditionally-Gaussian filter is of the following form:

$$
\begin{aligned}
d\bar{x}_t &= (A_1 \bar{x}_t + B_1)\, dt + S\, dv\,, \\
S &= (\Gamma_t\, C_1^{*})\, (\sigma\sigma^{*})^{-0.5}\,, \\
dv &= (\sigma\sigma^{*})^{-0.5}(dy_t - [C_1 \bar{x}_t + D_1\, dt])\,, \\
d\Gamma_t &= (A_1\, \Gamma_t + \Gamma_t\, A_1^{*} + G_1 G_1^{*} - SS^{*})\, dt\,,
\end{aligned}
\qquad (3\text{-}13)
$$

where A_1, B_1, C_1, D_1 are as in (3-12), and the arguments $(t,\bar{u}_t)$ are again omitted for brevity. A second-order example to demonstrate the above algorithm steps will be given in the following chapter.

Remarks:

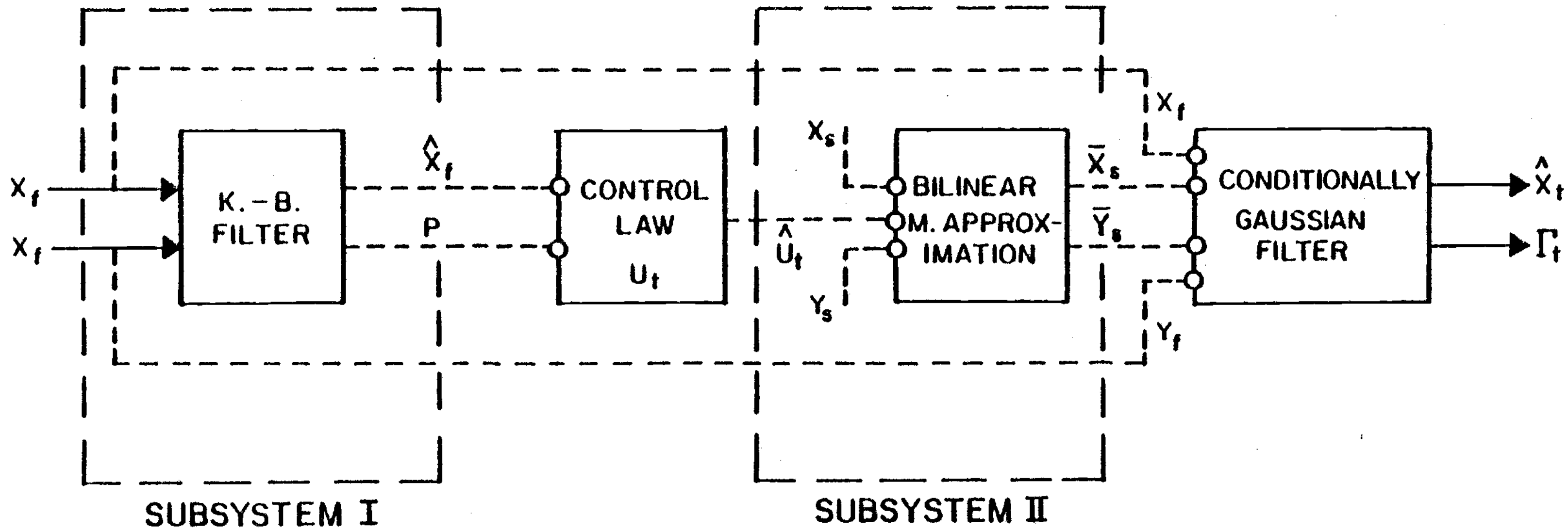
(i) The controls $\bar{u}_l$ $l=i=j=1,2,\ldots n+1$ are suboptimal, since they are functions of the observations in the first stage. If all of the observations are used in the first stage (i.e. the observations and the system are linear), then the controls

$\bar{u}_l$, $l = 1,2,\ldots n+1$ are optimal (in m.s.e).

(ii) In using this algorithm, the control parameters are obtained without solving the stochastic Bellman equation. However, it is still necessary to evaluate the conditional-expectation expressions in (3-9).

(iii) In general, the state estimate obtained in Step 5 is an improvement relative to that obtained in the first stage, and that due to the use of all the information in the last step.

(iv) If a "local-time" filtering criterion is assumed, then, as has been shown by Kolodziej and Mohler in [K3], there is no control-law calculation since the approximation parameters are not functions of u_t. But, the evaluation of similar conditional-expectation expressions are still required.



----o *Implies function or functional operator*

Figure 3-1

Block Diagram Representation of the TNF

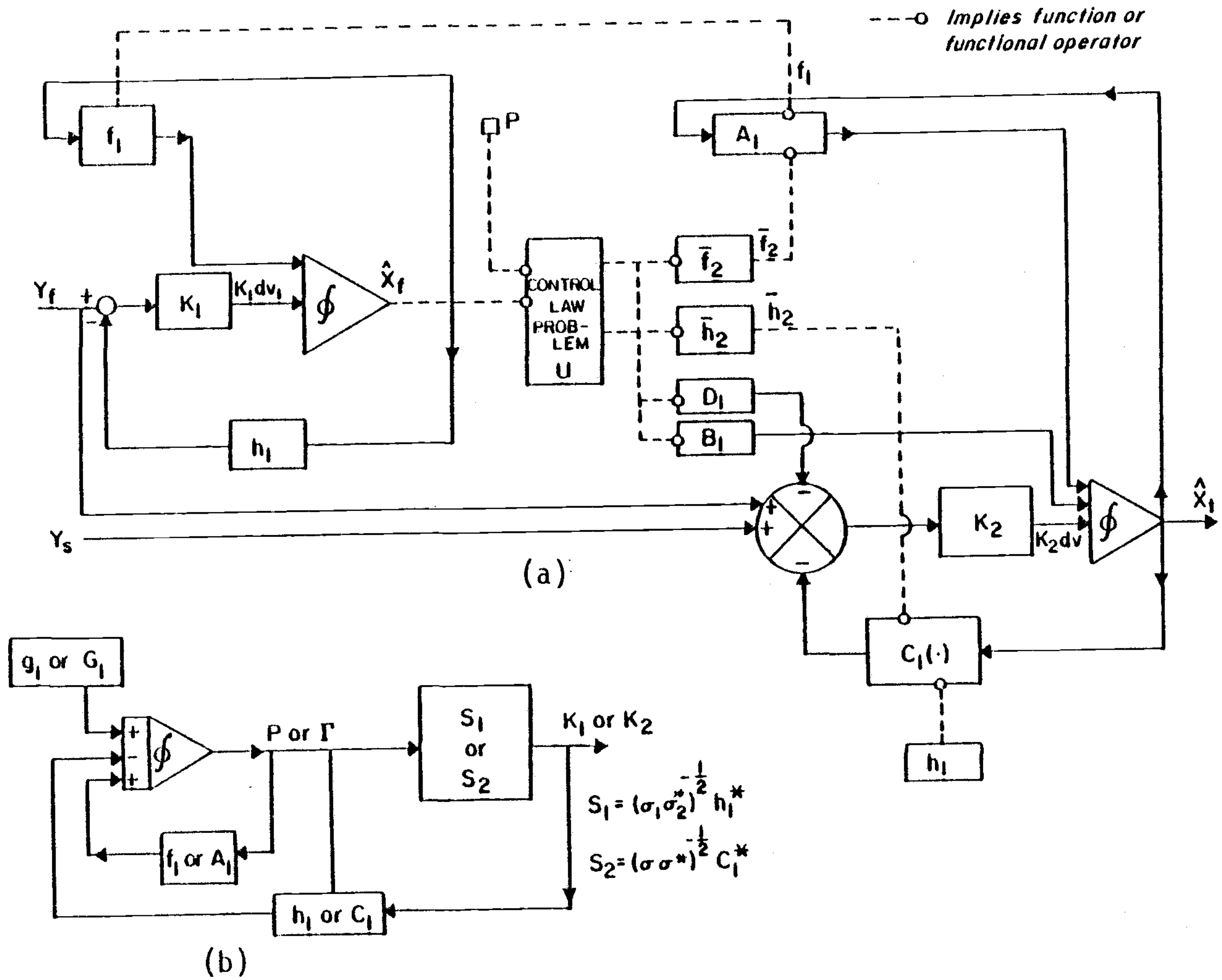


Figure 3-2
Schematic Representation for the TNF

Chapter 4

4. APPLICATIONS AND SIMULATIONS STUDIES

The implementation of the proposed two-step, nonlinear filter (TNF) are demonstrated via applications to a certain class of nonlinear system. In [K2] the applicability and effectiveness of the filter have been demonstrated for a nonlinear system that lacks the usually necessary model "smoothness". Obviously, the TNF is also applicable to "smooth" models as demonstrated herein.

The examples that are treated here were chosen from practical applications to illustrate the proposed procedures.

The claims made for improved performance, are verified through computer simulation results in the following sections.

4-1. General Simulation Comments

The digital-simulation examples that are discussed in the succeeding sections, were coded in FORTRAN in such a manner that the program generates solutions to the states of the original system, and the estimated values of the states by both the TNF and the EKF. Moreover, throughout all the simulation cases, the Wiener processes w_t which describe the excitation noises are generated from pseudo-random Gaussian variables v_i N(0,1). The latter generated by a standard (IMSL) Library Subroutine. And, increments of w_t are approximated by $dw \simeq \sqrt{\Delta t \, v_i}$, where Δt is the integration step-size (.001 to .0001). In the simulation, a fourth-order, Runge-Kutta integration algorithm is used for all trajectory filters and differential equations of both the original system and the error-covariance matrices.

The performance of the two filters are compared on the basis of the "mean-square error" (m.s.e) of the filter output to x_t such that:

$$J_{TNF} = \int_0^T (x(t) - \bar{x}_{TNF}(t))^2 dt \,,$$

$$J_{EKF} = \int_0^T (x(t) - \bar{x}_{EKF}(t))^2 dt \,, \tag{4-1}$$

and JJ gives the relative (percentage) difference between $J_{TNF}(t)$, $J_{EKF}(t)$ such that

$$JJ = \frac{J_{EKF}(t) - J_{TNF}(t)}{J_{EKF}} \times 100 \,. \tag{4-2}$$

Actually, numerous simulation tests were conducted during these simulation studies, however, only a small representative sampling of the results have been presented here. Nevertheless, these results, together with the ones presented in [K2], clearly demonstrate the effectiveness of the proposed filter. It should be noted that throughout the simulation process, the following identification symbols have been used to identify the different plots in the succeeding Figures of the sequel: the "□" (solid line) is used to identify the state trajectory of the original system, "∇" (dashed line) to mark the output of the TNF, while "x" (dashed line) is reserved for the output of the EKF.

4-2 Example 1

An important underwater application is the sonar tracking of a moving rigid body. This could be an active tracking of multi-mode range system or passive tracking with multi-receiver-transmitter and correlated time delay. The rigid body considered here is a point mass.

4-2-a Problem Definition and Model Formulation

The problem at hand is to develop an optimal, (or at least suboptimal), nonlinear finite-dimensional estimation algorithm for the range and range rate of a rigid body based on the noisy observation of its position and velocity provided by a sonar signal.

The state vector, (Range = x_1, Range Rate = x_2), evolves according to the following stochastic differential equation [P1]:

$$dx = A\,x\,dt + Gdw \quad , \tag{4-3}$$

where, $dx = \begin{pmatrix} dx_1 \\ dx_2 \end{pmatrix}$, $A = \begin{pmatrix} 0 & 1 \\ 0 & -\alpha \end{pmatrix}$, $G = \begin{pmatrix} 0 \\ \alpha \end{pmatrix}$, $\alpha = 1/\tau$, w is a Wiener process. Here x_2 is influenced randomly by target maneuver which is characterized by τ the maneuver time constant.

The measurement equations are nonlinear due to the "acoustic propagation" time delay, when tracking is derived directly in current time. The observation equations are

$$dy = H(x,t)dt + Rd\, w_2 , \tag{4-4}$$

where $dy = \begin{pmatrix} dy_1 \\ dy_2 \end{pmatrix}$, $H(x,t) = \begin{pmatrix} x_1 + B_1 x_1 x_2 \\ x_2 \end{pmatrix}$, $R = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix}$,

w_2 is a Wiener vector of measurement noises, $B_1 = a/c$, a is a constant, and c is the average speed of sound in water.

4-2-b Filtering Algorithm

To find an approximate finite-dimensional filter for the nonlinear system in (4-3), (4-4), the decomposition scheme and the corresponding filtering algorithm, discussed in the previous chapter, are utilized.

Subsystem I

Here only the range rate information is used. Thus, the linear system and observation equations are:

$$dx_1 = x_2 dt ,$$

$$dx_2 = - \alpha \; x_2 + c_1 \; \alpha \; dw_1 \; ,$$

$$dy_2 = x_2 \; dt + \sigma_2 \; dw_2 \; , \tag{4-5}$$

where c_1 is a multiplicative constant. Using the Kalman-Bucy method, the estimate $\bar{x}_{if}$, $i = 1,2$ satisfies the following linear stochastic equations:

$$d\bar{x}_{1f} = \bar{x}_{2f} dt + \frac{P_3}{\sigma_2} \, dv \quad ,$$

$$d\bar{x}_{2f} = -\alpha \bar{x}_{2f} dt + \frac{P_2}{\sigma_2} \, dv,$$

where $\bar{x}_{if} = E(xi/y_2)$, $i = 1, 2$, the conditional expection, and dv is the innovation process. The corresponding error-covariance equations are:

$$dP_1 = (2P_3 - P_3{}^2/ \, \sigma_2{}^2) \; dt, \tag{4-6}$$

$$dP_2 + (-2\alpha P_2 + c_1^{\,2}\alpha^2 - \frac{P_2^{\,2}}{\sigma_2^{\,2}})\ dt\ ,$$

$$dP_3 = (P_2 - \alpha P_3 - \frac{P_2P_3}{\sigma_2^{\,2}})\ dt\ .$$

Subsystem II

$$\text{Let}\quad x_1x_2 \simeq (u_1x_1 + u_2x_2 + u_3)\ . \tag{4-7}$$

Here u_i, $i = 1,2,3$ are measurable with respect to the σ-algebra $(y_{2s};\ s\varepsilon[0,t])$, and chosen to minimize the following global criterion:

$$J(u) = \min_{u_i} E\ (\int_0^T (x_1x_2 - (u_1x_1 + u_2\ x_2 + u_3))^2)dt. \tag{4-8}$$

Equation (4-8) can be written equivalently using the properties of expectation as

$$J(u) = \min_{u_i} E\ [\int_0^T E\ (x_1x_2 - (u_1x_1 + u_2x_2 + u_3))^2/y_{2s};s\varepsilon[0,T])\ dt]$$

$$= \min_{u_i} E\ [\int_0^T E^t(x_1^{\,2}x_2^{\,2} - R + K^2)\ dt]\ , \tag{4-9}$$

where $\quad K = u_1x_1 + u_2x_2 + u_3x_3\ ,$

$$R = 2\, x_1 x_2 K \quad .$$

Let $H(x,u) = E^t(x_1{}^2 x_2{}^2) - E^t(R) + E^t(K^2)$. (4-10)

Then (4-9) becomes

$$J(u) = \min_{u_i} E\left(\int_0^T H(x,u)\, dt \right) \quad . \tag{4-11}$$

Here, the minimization in (4-41) with respect to u_i, $i = 1,2,3$ can be performed since the expectation is conditioned on the σ-algebra $(y_{2s};\ s\varepsilon[0,t])$ which is not a function of the parameters u_i, $i = 1,2,3$. And the fact that the unobservable states x_i, $i = 1,2$ are not functions of the control parameters u_i, $i = 1,2,3$. Accordingly, $\dfrac{\partial E^t(x_1{}^2 x_2{}^2)}{\partial u_i} = 0$ for $i = 1,2,3$. Now, using similar arguments as in Chapter 3, the minimization of (4-11) requires that

$$V_t + 0.5\ \mathrm{tr}\,(\sigma\, V_{xx}\, \sigma^*) + \min_{u_i}\, [V_x \quad x_1 x_2 + H(x,u)] = 0 \text{ for } i=1,2,3 \ . \tag{4-12}$$

Thus, performing the minimization with respect to u_i, $i=1,2,3$ in (4-12), the following set of equations (which must be solved simultaneously) are obtained:

$$- 2\, E^t(x_1{}^2 x_2) + 2\bar{u}_1\, E^t\, (x_1{}^2) + 2\bar{u}_3\, E^t(x_1) + 2\bar{u}_2 E^t(x_1 x_2) = 0,$$

$$- 2\, E^t(x_2{}^2 x_1) + 2\, \bar{u}_2\, E^t(x_2{}^2) + 2\, \bar{u}_3\, E^t(x_2) + 2\, \bar{u}_1 E^t(x_1 x_2) = 0\ ,$$

$$-2\, E^t(x_1 x_2) + 2\, \bar{u}_3 + 2\, \bar{u}_1\, E^t(x_1) + 2\, \bar{u}_2 E^t(x_2) = 0\ . \tag{4-13}$$

To calculate $\bar{u}_1$, $\bar{u}_2$, $\bar{u}_3$, which minimize (4-9) or (4-11), from (4-13) the conditional expectation expressions must be evaluated first. From the basic definition,

$$E^t(x_i{}^2) = P_i + \bar{x}_i{}^2, \quad i = 1,2\ . \tag{4-14}$$

To evaluate the expressions $E^t(x_i{}^2 x_j)$, $E^t(x_i x_j{}^2)$, $E^t(x_i x_j)$, $i,j = 1,2$, the results of Lemmas (B1, B2) from Appendix B are utilized. Thus,

$$E^t(x_i{}^2 x_j) = 2\bar{x}_i u_{ij} + \bar{x}_j\, u_{ii}\ ,$$

$$E^t\, (x_i x_j{}^2) = 2\bar{x}_j\, u_{ij} + \bar{x}_i\, u_{jj}\ , \tag{4-15}$$

$$E^t(x_i x_j) = \bar{x}_i \bar{x}_j + P_3\ .$$

Then, from (4-14), (4-15) in (4-13),

$$\bar{u}_1 = \bar{x}_{2f} , \qquad (4\text{-}16)$$

$$\bar{u}_2 = \bar{x}_{1f} ,$$

$$\bar{u}_3 = P_3 - \bar{x}_{1f}\bar{x}_{2f} .$$

Hence, $x_1 x_2 \approx \bar{x}_{2f} x_1 + \bar{x}_{2f} x_2 + P_3 - \bar{x}_{1f}\bar{x}_{2f}$. (4-17)

Then, the new equivalent system is

$$dx \approx A\, x\, dt + G dw_1 ,$$

$$dy \approx \bar{H}\, x\, dt + D(\bar{x})\, dt + \bar{R} dw_2 , \qquad (4\text{-}18)$$

where $\bar{H} = \begin{bmatrix} B_1(1 + \bar{u}_1) & \bar{u}_2 \\ 0 & 1 \end{bmatrix}$, $D(\bar{x}) = \begin{bmatrix} B_1\, \bar{u}_3 \\ 0 \end{bmatrix}$,

$\bar{R} = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}$; A, G are the same as in (4-3).

Now, assume the following:

(i) if f denotes any of the functions A, G, $\bar{H}$, D, $\bar{R}$, then

$$P\left(\int_0^T | f | \, dt < \infty \right) = 1;$$

(ii) x_0 given y_0 is conditionally Gaussian. Then, from (2-12) the corresponding conditionally-Gaussian filter is

$$d\bar{x}_1 = \bar{x}_2 dt + \frac{1}{\sigma_1^2} \left(\Gamma_1 \, (1 + B_1\bar{u}_1) + B_1\bar{u}_2 \, \Gamma_3\right) \quad dv_1 + \frac{\Gamma_3}{\sigma_2^2} \, dv_2 \ , \tag{4-19}$$

$$d\bar{x}_2 = -\alpha x_2 \, dt + \left[\frac{1}{\sigma_1^2} \left(\Gamma_3(1 + B_1\bar{u}_1) + B_1\bar{u}_2\Gamma_2\right)\right] dv_1 + \frac{\Gamma_2}{\sigma_2^2} \, dv_2 \ .$$

where,

$$dv_1 = dy_1 - (1 + \bar{u}_1 B_1 \bar{x}_1 + B_1\bar{u}_2\bar{x}_2 + B_1\bar{u}_3)dt \ ,$$

$$dv_2 = dy_2 - \bar{x}_2 dt \ ,$$

and B_1,α are defined as before.

The covariance equations are

$$d\Gamma_1 = \left\{ 2\Gamma_3 - \left[\frac{1}{\sigma_1^2} \left(\Gamma_1(1 + B_1\bar{u}_2) + \Gamma_3 B_1\bar{u}_2\right)^2 + \frac{\Gamma_3^2}{\sigma_2^2} \right]\right\} dt \ ,$$

$$d\,\Gamma_2 = \left\{ c_1^2\alpha^2 - 2\alpha \, \Gamma_2 - \left[\frac{1}{\sigma_1^2} \left(\Gamma_3(1 + B_1\bar{u}_1) + \Gamma_2 B_1\bar{u}_2\right)^2 + \frac{\Gamma_2^2}{\sigma_2^2} \right]\right\} dt,$$

$$d\,\Gamma_3 = \{\Gamma_2 - \alpha\Gamma_3 - [\frac{1}{\sigma_1^2}(\Gamma_1(1+B_1\bar{u}_1) + \Gamma_3 B_1\bar{u}_2)\,(\Gamma_3(1+B_1\bar{u}_1)$$

$$+ \Gamma_2 B_1 \bar{u}_2)] - \frac{1}{\sigma_2^2}(\Gamma_2\Gamma_3)\}\,dt, \qquad (4\text{-}20)$$

where $\bar{u}_1$, $\bar{u}_2$, $\bar{u}_3$ are as in (4-16).

Notice that in this case (4-19), (4-20) are the same as the filter equations of the modified-second-order truncated filter defined by Jazwinski [JI] because the nonlinearity is of second order, although the approach is different.

4-2-c Extended Kalman Filter

The filter equations are [J1]:

$$d\,\bar{x}_1 = \bar{x}_2 dt + \frac{1}{\sigma_1^2}[(1+B_1\bar{x}_2)\,p_1 + p_3 B_1 \bar{x}_1]\,dv_1$$

$$+ \frac{p_3}{\sigma_2^2}\,dv_2\,, \qquad (4\text{-}21)$$

$$d\bar{x}_2 = -\alpha\ \bar{x}_t dt + \frac{1}{\sigma_1^{\,2}} [(1 + B_1\bar{x}_2)\ P_3 + P_2 B_1 \bar{x}_1)]\ dv_1$$

$$+ P_2 / \sigma_2^{\,2}\ dv_2\ ,$$

where B_1, α are as defined before, and

$$dv_1 = (dy_1 - (1 + B_1\bar{x}_2)\ \bar{x}_1 dt),$$

$$dv_2 = (dy_2 - \bar{x}_2 dt)\ .$$

The covariance equations are:

$$dP_1 = \{ 2P_3 - \frac{1}{\sigma_1^{\,2}} ((1 + \bar{x}_2 B_1)\ P_1 + \bar{x}_1 B_1 P_3)^2 - \frac{P_3^{\,2}}{\sigma_2^{\,2}} \}\ ,$$

(4-22)

$$dP_2 = \{ c_1^{\,2}\alpha^2 - 2\alpha P_2 - \frac{1}{\sigma_1^{\,2}} [(1 + B_1\bar{x}_2)P_3 + B_1\bar{x}_1 P_2]^2 - \frac{P_2^{\,2}}{\sigma_2^{\,2}} \}\ dt\ ,$$

$$dP_3 = \{ P_2 - \alpha P_3 - [\frac{1}{\sigma_1^{\,2}} ((1 + B_1\bar{x}_2)P_1 + \bar{x}_1 B_1 P_3)\ ((1 + B_1\bar{x}_2)$$

$$P_3 + B_1 P_2 \bar{x}_1) - \frac{1}{\sigma_2^{\,2}} P_2 P_3 \}\ dt\ .$$

4-2-d Simulation Results

Here, the previous sonar problem is simulated, where both equations (4-3), (4-4), and the TNF, EKF equations, (4-19), (4-20), (4-2i), (4-22), were solved by the digital computer. Numerous simulation tests were conducted for this problem using different initial conditions and parameter values. But only a representative sampling of the results have been presented here. These simulation results were tabulated in Tables (1,2), where JJ_1 represents the percentage position (range) error (in m.s.e) accuracy of the TNF as compared to the EKF, while JJ_2 represents the percentage velocity (range rate) error (in m.s.e) accuracy of the TNF as compared to the EKF range-rate output. In all cases, JJ_1, JJ_2 were calculated according to the equations in (4-1). In figures (4-9) - (4-16) the range root-mean-square (rms) error $Q_1(t)$, $Q_2(t)$, and the rms velocity error $VQ_1(t)$, $VQ_2(t)$, for both the TNF, the EKF respectively, were calculated as follows:

$$Q_j(t) = \left[\sum_{i=1}^{IL} \frac{(x_j^{(i)}(t_k) - \bar{x}_j^{(i)}(t_k))^2}{IL} \right]^{\frac{1}{2}} , \tag{4.23}$$

$$VQ_j(t) = \left[\sum_{i=1}^{IL} \frac{(x_j^{(i)}(t_k) - \bar{x}_j^{(i)}(t_k))^2}{IL} \right]^{\frac{1}{2}} , \tag{4-24}$$

where $(x_j^{(i)}(t_k), \bar{x}_j^{(i)}(t_k))$ are the j^{th} components of the true state and its corresponding TNF, EKF estimates at time t_k on the i^{th} simulation run, in a series of IL runs. For completeness, some comments on the filter initialization seem in order here. Under actual operating conditions it is extremely difficult, and indeed rare due to one reason or another, to obtain reliable initial estimates of the state vector and its associated covariance matrix. Consequently, the following set of initial conditions are realistically chosen. Throughout, the initial range value is 5000 meters, while the initial range rate value is assumed constant and chosen from the following set. (50m/sec, 500 m/sec, 1000 m/sec). The initial condition of the estimates are calculated according to the following equation:

$$\bar{x}_i(0) = x_i(0) + \sqrt{p_i(0)}\,\eta_i \quad , \; i=1,2 \tag{4-25}$$

where η_i is a random noise. The initial covariance matrix is:

$$P(0) = \begin{bmatrix} P_1(0) & P_3(0) \\ P_3(0) & P_2(0) \end{bmatrix} = \begin{bmatrix} 10^6 & 10^2 \\ 10^2 & 10^4 \end{bmatrix} ,$$

where the diagonal elements of P(0) are chosen relatively large so that the filter will "forget" the initial values as more data arrived, and to ensure the randomness of the intial estimates. In all cases, a system noise of 1% of the initial state values is used, and different levels of measurement noises (from 2% - 20%) of the initial range, range rate respectively, are added. For convenience, the time interval for each run is 10 seconds, and the number of runs for each simulation test case is between 10 to 20 runs. Thus, all results have been ensemble average over IL runs; [the number of runs for each simulation test (10 to 20)].

The effect of increasing the nonlinearity, (i.e. increases a), of the system on the filter trajectories and the rms error levels are demonstrated in Figures (4-2), (4-6), (4-10), and (4-14) as compared to Figures (4-1), (4-5), (4-9), and (4-13) respectively. Accordingly, the TNF performance improved substantially, and the rms-error levels increased enormously as compared to the rms-error levels of TNF. The above conclusions are also demonstrated by Table (1).

In comparing Figures (4-3), (4-7), (4-11), and (4-15) with Figures (4-2), (4-6), (4-10), and (4-14), it is noticed that an increase in the velocity measurement noise standard deviation σ_2 by 10% will degrade the performance of the TNF, and improve the performance of the EKF, while the rms-error levels significantly decrease. However, the TNF still performs better than the EKF.

Comparison of Figures (4-4), (4-8), (4-12), and (4-16) with Figures (4-1), (4-5), (4-9), and (4-13) respectively, and Table (2) indicates that the EKF gains in accuracy relative to the TNF as the observations become more noisy (i.e. increases the range measurement noise standard deviation, σ_1 to 20%). This is due to the fact that the nonlinearity, (here in the range measurement equation), is masked by the large measurement noise. In essence, this is in total agreement with the remarks pointed out by Jazwinski [J1] in his criticism of Schwartz's simulations [S2].

Finally, it is generally noted that the mean-square errors $Q_1^2(t)$, $VQ_1^2(t)$, $Q_2^2(t)$, $VQ_2^2(t)$ of the estimators and the optimal error covariance $\Gamma_1(t)$, $\Gamma_2(t)$, and $P_1(t)$, $P_2(t)$ of the TNF, EKF respectively are not the same. That is due to the fact that averaging over (IL) samples paths (10-20) does not give a good approximation to the expectation.

From the tables and figures mentioned above, the following remarks seem in order:

(i) In almost all the cases, (except Table 2), the TNF shows significant improvement in filter accuracy as compared to the EKF.

(ii) In general, the TNF range rms error, and velocity rms error are much smaller than corresponding EKF as shown in Figures (4-13) - (4-16). This clearly demonstrates the effectiveness of TNF over the EKF.

(iii) The computer utilization cost of the TNF is around 10% higher than the cost of the EKF, while the storage requirements are relatively equal. Thus, the computer cost consideration will not be a preference factor as far as this application is concerned.

(iv) Finally, the complexity of the proposed algorithm (TNF) over the EKF might be justified by the significant improvement in the filter accuracy.

4.3 Example 2

In this example a new nonlinear filtering and tracking technique [K3], which is similar to the TNF, is applied to a passive-sonar-two dimensional problem. This technique is suited for the type of problem under consideration, where the observation equation is scalar. Thus, application of the TNF will encounter difficulties with respect to control policy calculation. However, this new method does not require control calculation, but unfortunately, the global properties are lost.

4-1-a. Problem Definition and Model Formulation

This problem describes the two-dimensional, bearings-only target motion [A2]. Figure (4-17) presents a geometric configuration of both the target and the observer, where it is assumed that both lie in the same horizontal plane.

It is assumed that the system behavior evolves according to the following stochastic differential equations:

$$\frac{dx}{dt} = A\, x_t + B\, u_t + G\, w_{1t}\ , \qquad (4\text{-}26)$$

where

$$x = \begin{bmatrix} x_1, \text{ relative range component in the x direction} \\ x_2, \text{ relative range component in the y direction} \\ x_3, \text{ relative velocity component in the x direction} \\ x_4, \text{ relative velocity component in the y direction} \end{bmatrix},$$

$$u = \begin{bmatrix} u_x - u_{ox}, \text{ relative acceleration in the x direction} \\ \\ u_y - u_{oy}, \text{ relative acceleration in the y direction} \end{bmatrix},$$

u_x, u_y are acceleration of the target in the (x,y) direction respectively,

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ -1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad G = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ R_1 & 0 \\ 0 & R_2 \end{bmatrix}.$$

Here w_{1t}, (w_{1tx}, w_{1ty}), is an additive noise to model random target-acceleration fluctuations from the assumed constant velocity trajectory, (with $u_x = u_y = 0$).

The measurement data consists entirely of passive sonar bearings as follows:

$$\beta(t) = \text{Arctan}\,[x_1(t)/x_2(t)] + \sigma n(t)\ , \tag{4-27}$$

where $\beta(t)$ is the measured target bearings, and $\eta(t)$ is an independent additive Gaussian measurement noise and has variance $\sigma^2(t)$. Moreover, the observation is almost spacially continuous if we consider the scenario for an isolated submarine tracking a surface ship or far distance submarine. Now, if we let $dy = \beta dt$ then (4-27) can be written as

$$dy = H(x_1,x_2,t)\ dt + \sigma\ dw_2\ , \tag{4-28}$$

where $dw_2 = \eta_t\ dt$ is a Wiener process, $H(t) = \tan^{-1}(\frac{x_1(t)}{x_2(t)})$.

4-3-b. Filtering Algorithm

The system in (4-26), (4-28) is a nonlinear system with nonlinear observations. In [A2], [C1], it has been shown that the EKF suffers from the "ill-conditioning" phenomena due to the error covariance-matrices false observability [A2]. The new approach [K3] proceeds as follows: First, the approximating system is defined with

$$\bar{H}\ (x_1,x_2,y,t) = h_0(y) + [h_1(y),\ h_2(y)]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \tag{4-29}$$

Now assume the following:

(i) if f denotes any of the functions F, B, G, h_0, h_1, h_2, then

$$P\left(\int_0^T |f|\, dt < \infty\right) = 1;$$

ii) x_0 given y_0 is conditionally Gaussian;

iii) h_0, h_1, h_2 are y_t measurable.

Then, an approximating model for the target and receiver equations (4-26), (4-28) is presented by:

$$dx \simeq (Ax + Bu_0)\, dt + G\, dw_1 \,,$$
$$dy \simeq (h_0 + kx)\, dt + \sigma\, d\, w_2 \,, \tag{4-30}$$

where $k = (h_1, h_2)$, w_1, w_2 are independent Wiener processes, $u_0 = (u_{0x}, u_{0y})$, the observer acceleration. The arguments (x_1, x_2, y, t), (y) are omitted. Also, the assumptions significantly provide that in (4-30) x_t, y_t are conditionally Gaussian. Hence, from (2-12), (see Chapter 2-Section 2-3), the recursive formulas for $\bar{x}_t$ (the conditional mean), and Γ_t (the conditional covariance) are:

$$d\bar{x}_t = (A\bar{x} + Bu_0)\, dt + \frac{S}{\sigma}\,[dy - (h_0 + k\bar{x}_t)\, dt] \,,$$
$$d\Gamma_t = (A\Gamma_t + \Gamma_t A^* + GG^* - SS^*)\, dt \,, \tag{4-31}$$

where $S = \dfrac{\Gamma_t M^*}{\sigma}$, $M = [h_1, h_2, 0, 0]$.

The recursive formulas in (4-31) will completely characterize the filter equations if h_0, h_1, h_2 are obtained. Now, using the following mean-square criterion:

$$J(u) = E(||F - \bar{F}||^2) . \tag{4-32}$$

Now, from (4-28), (4-30) equation (4-32) can be written as

$$J(u) = E\{(H - (h_0 + kx)(H - (h_0 + x^* k^*)\}$$

$$= E\{E^t(H^2) + h_0{}^2 + kE^t (xx^*) k^* - 2h_0 E^t(H)$$

$$+ h_0 kE^t(x) + h_0E^t(x^*) k^* - E^t(Hx^*) k^* - KE^t(Hx)\} , \tag{4-33}$$

where again E^t ( ) is the conditional expectation operator.

Let

$$a = E^t(xx^*) p^{-1},$$

$$b = p^{-1},$$

$$c = E^t(x^*) p^{-1}, \tag{4-34}$$

$$d = E^t(H^*)-(E^t(H))^2- \ E^t(x^*)\ E^t(Hx^*)\ E^t(H) + E^t(x)\ E^t(Hx)\ E^t(H)$$

$$-E^t(Hx^*)\ E^t(Hx) - E^t(x)\ E^t(x^*)[E^t(H)]^2 \ \ P^{-1},$$

$$e_0 = h_0 + k\ E^t(x) - E^t(H)\ ,$$

$$e_1 = h_0\ E^t(x^*) + k\ E^t(xx^*) - E^t(Hx^*).$$

Here P is a 2x2 positive definite matrix.

Now, (4-33) can be written, using (4-34), as

$$E\ \{(a\ e_0e_0^* + b\ e_1e_1^* - c\ e_0e_1^* - e_1e_0^*c^* + d)\}\ . \tag{4-35}$$

Since ab > c^*c, then the minimization of (4-35) requires that $e_0 = 0$, $e_1 = 0$. Thus,

$$h_0 = E^t(H) - [E^t(Hx^*) - E^t(H)E(x^*)]\ P^{-1}\ E^t(x)\ ,$$

(4-36)

$$k = [E^t(Hx^*) - E^t(H)\ E^t(x^*)]\ p^{-1} = [h_1,h_2]\ .$$

Hence, from (4-36), (4-29) becomes

$$\bar{H}\ (x_1,x_2,y,t) = E^t(H) + [E^t(Hx^*) - E^t(H)\ E^t(x^*)]p^{-1}(x-E^t(x))\ .$$

(4-37)

To evaluate the conditional-expectation terms in (3-36), the following approximation, (see (2-13), (2-14)), can be used.

Let f denote either of the functions (H, Hx^{*}), then,

$$E^{t}(f) \simeq \int_{R} f(\xi, y_t, t)\, d\,\Phi\,(\bar{x}, \Gamma_t, \xi) \equiv \bar{f}\,(\bar{x}, y, \Gamma, t)\ , \tag{4-38}$$

where $d\,\Phi\,(\bar{x}_t, \Gamma, \xi)$ is a differential Gaussian measure as defined in (2-14). Thus, using (4-38), (2-14)

$$E^{t}(H) \simeq \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \arctan\left(\frac{x_1}{x_2}\right) \frac{1}{2\pi\,|\Gamma_t|}\, \mathrm{Exp} - 0.5\,[(x-\bar{x})^{*}\Gamma^{-1}(x-\bar{x})]\, dx_1 dx_2\ , \tag{4-39}$$

$$E^{t}(Hx^{*}) = [E^{t}(Hx_1),\ E^{t}(Hx_2)]\ , \tag{4-40}$$

$$E^{t}(Hx_1) = \frac{1}{2\pi\ |\Gamma_t|} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 \arctan \frac{x_1}{x_2}\, \mathrm{Exp} - 0.5\,[(x-\bar{x})^{*}\Gamma^{-1}\,(x-\bar{x})]\, dx_1 dx_2, \tag{4-41}$$

$$E^{t}(Hx_2) = \frac{1}{2\pi\ |\Gamma_t|} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_2\ \arctan\left(\frac{x_1}{x_2}\right) \mathrm{Exp} - 0.5\,[(x-\bar{x})^{*}\ \Gamma^{-1}(x-\bar{x})]\, dx_1 dx_2.$$

But, analytical evaluation of (4-39), (4-41) is very difficult. Thus, the following approximation scheme is used, where the arctan function is first expanded by Taylor series for a function of two variables, then find the conditional expectation for each terms using equation (4-38), (2-14). Thus, using the Taylor series expansion up to the fourth term (n=4), and utilizing (4-38), (2-14), Lemma B2 (Appendix B), (4-39) becomes

$$E^t[\arctan \frac{x_1}{x_2}] \approx \arctan(\frac{\bar{x}_1}{\bar{x}_2}) + \frac{(\bar{x}_1\bar{x}_2(\Gamma_2-\Gamma_1) + \Gamma_{12}(\bar{x}_1^{\,2}-\bar{x}_2^{\,2})}{a^2}$$

$$+ [3\Gamma_2(\bar{x}_1^{\,4} + \bar{x}_2^{\,4} - 6\,\bar{x}_1^{\,2}\,\bar{x}_2^{\,2})(\Gamma_1-\Gamma_2) \tag{4-42}$$

$$+ \bar{x}_1\bar{x}_2(\bar{x}_2^{\,2}-\bar{x}_1^{\,2})(3\Gamma_1^{\,2}+3\Gamma_2^{\,2} - 6(\Gamma_1\Gamma_2 + 2\Gamma_{12}^{\,2}))]/a^4 \quad ,$$

where $a = (\bar{x}_1^{\,2} + \bar{x}_2^{\,2})$.

Also (4-41) becomes

$$E^t(x_1\arctan \frac{x_1}{x_2}) \approx \bar{x}_1 E^t(\arctan \frac{x_1}{x_2}) + A_1 + B_1 + C_2\Gamma_1^{\,2}\,\bar{x}_2 \tag{4-43-a}$$

$$+ C_1\,\Gamma_1\,\Gamma_{12}\,\bar{x}_1 + C_3\,\bar{x}_1^{\,2}\,\bar{x}_2 \;,$$

$$E^t(x_2 \arctan \frac{x_1}{x_2}) \simeq \bar{x}_2 E^t(\arctan \frac{x_1}{x_2}) + A_2 + B_1 + C_1 \Gamma_2{}^2 \bar{x}_1 \tag{4-43-b}$$

$$+ C_2 \Gamma_2 \Gamma_{12} \bar{x}_2 + C_3 \bar{x}_2{}^2 \bar{x}_1 ,$$

where

$$A_1 = \frac{[\Gamma_1 \bar{x}_2 - \Gamma_{12} \bar{x}_1]}{a} , \qquad A_2 = \frac{[\Gamma_{12} \bar{x}_2 - \Gamma_2 \bar{x}_1]}{a} , \tag{4-43-c}$$

$$B_1 = \frac{(\Gamma_1 \Gamma_2 + 2\Gamma_{12}{}^2)(\bar{x}_2{}^3 - \bar{x}_1{}^3)}{a^3} ,$$

$$C_1 = \frac{(\bar{x}_1{}^2 - 3\bar{x}_2{}^2)}{a^3} , \qquad C_2 = \frac{(3\bar{x}_1{}^2 - x_2{}^2)}{a^3} , \tag{4-43}$$

$$C_3 = -3 \Gamma_1 \Gamma_3 \bar{x}_1 C_1 ,$$

$$C_4 = -3 \Gamma_2 \Gamma_3 \bar{x}_2 C_2 .$$

Therefore, from (4-42) and (4-43), (4-36) becomes

$$h_1 = d_1[k_1 \Gamma_2 - k_2 \Gamma_{12}] , \tag{4-44}$$

$$h_2 = d_1 \, [k_2 \, \Gamma_1 - k_1 \, \Gamma_{12}],$$

where

$$k_1 = [A_1 + B_1 + C_1 \, \bar{x}_1 \, \Gamma_1 \, \Gamma_3 + C_3 + C_2 \, \bar{x}_2 \, \Gamma_1{}^2] \; ,$$

$$k_2 = [A_2 + B_1 + C_2 \, \bar{x}_2 \, \Gamma_1 \, \Gamma_3 + C_4 + C_2 \, \bar{x}_1 \, \Gamma_2{}^2] \; ,$$

$$d_1 = [1/(\Gamma_1 \, \Gamma_2 - \Gamma_{12}{}^2)] \; ,$$

A_1, B_1, C_1, C_2, C_3 are defined in (4-43). But, in (4-31)

$$S = \frac{1}{\sigma} \, [\Gamma_t \, M^t] \; , \; M = [h_1, \, h_2].$$

Then, from (4-44),

$$S = \frac{1}{\sigma} \begin{bmatrix} \Gamma_1 \, h_1 + \Gamma_{12} \, h_2 \\ \Gamma_{21} h_1 + \Gamma_2 \, h_2 \\ \Gamma_{13} h_1 + \Gamma_{23} \, h_2 \\ \Gamma_{14} h_1 + \Gamma_{24} \, h_2 \end{bmatrix} = \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{bmatrix} . \tag{4-45}$$

Thus, from (4-45), the filter equations (4-31) become

$$d\bar{x}_t = (A \, \bar{x}_t + Bu_o) \, dt + \frac{S}{\sigma} \left(dy - E^t(H) \, dt \right) \; ,$$

$$d\,\Gamma_t = (A\,\Gamma_t + \Gamma_t\,A^* + GG^* - SS^*)\,dt\,, \tag{4-46}$$

where S is given by (4-45), $E^t(H)$ by (4-42), and A, G, B by (4-26). Here Γ_t is a 4x4 matrix, while $\bar{x}$ is a 4x1 vector.

4-3-c

The corresponding EKF equations are

$$d\bar{x}_t = (A\,\bar{x}_t + Bu_o)\,dt + \frac{P_t\,N^*}{\sigma^2}\,(dy\text{-}\arctan\frac{\bar{x}_1}{\bar{x}_2}\,dt)\,,$$

$$d\,P_t = (A\,P_t + P_t\,A^* + GG^* - \frac{P_t\,N^*\,N\,P_t}{\sigma^2})\,dt\,, \tag{4-47}$$

where $N = \left[\frac{\partial\,H_t\,(\bar{x},t)}{\partial x_i}\right] = \begin{bmatrix} \bar{x}_2/\,(\bar{x}_1{}^2 + \bar{x}_2{}^2)^2 \\ -\bar{x}_1{}^2\,/(\bar{x}_1{}^2 + \bar{x}_2{}^2)^2 \\ 0 \\ 0 \end{bmatrix}$,

and (A, B, G, σ) are as defined in (4-26).

4-3-d Simulation Results

The passive sonar problem described by equation (4-26), (4-28) and the estimated algorithm by both TNF, EKF which have been described by (4-46), (4-47) respectively, are simulated on the digital computer.

To compare the estimator's performance, the following scenario is devised. The target is at an initial range of 2700 yards, moves at a constant speed of 675.13 yards/min., and maintains a steady course of 0^o. In addition, the initial bearing is 0^o. Own-ship is assumed to be at the origin initially, maintains a constant speed of 954.63 yards/min., but periodically executed 90^o course changes as follows:

$$\text{from } 45^o \text{ to } -45^o \text{ at } t = (4 + 17k) \; [k = 0,1] \quad ,$$

$$\text{from } -45^o \text{ to } 45^o \text{ at } t = (12.5 + 17k) \; [k = 0,1].$$

The own-ship course changes at the rate of 3^o/second.

Numerous simulation tests were conducted for this problem using different levels of measurement noises and initial conditions. But only representative sampling of the results have been presented here. These results are tabulated in Table 3 and are shown in Figures (4-18) - (4-25) for additive, rms, measurement-noise levels of 3^o, 12^o, respectively, and JJ_1, JJ_2 are as defined in (4-22).

The initial state values were respectively

$$x(0) = [0,\ 2700,\ -\ 675.13,\ 0]\ ,$$

while the estimates were initialized according to the following equation

$$\bar{x}_i(0) = x_i(0) + \sqrt{p_i(0)}\ \eta_i,\ i = 1,2,3,4\ ,$$

where n_i is a random noise. The initial covariance matrix was

$$p(0) = \text{diag}\ [10^6,\ 10^6,\ 50^2,\ 50^2]\ .$$

Moreover, in all the simulation cases, system noises are added to the velocity states (x_3, x_4) to compensate for random target-acceleration fluctuations from the assumed constant-velocity trajectory ($u_{xt} = u_{yt} = 0$).

For convenience, the time interval for each run is 20 minutes, and the number of runs for each simulation test case is 10 runs. Thus, all results have been ensemble averaged over 10 runs.

The relative range trajectory and its corresponding estimates by both the TNF, EKF were simultaneously plotted in Figures (4-18) - (4-19) for the rms-noise levels of 3°, 12° respectively. And the relative-velocity trajectory and its estimates were shown in Figures (4-20), (4-21) respectively. Figures (4-22)-(4-23) showed the range rms-estimation errors ($Q_1(t)$, $Q_2(t)$) associated with TNF, EKF respectively and calculated in accordance to equation (4-23) for the same noise levels as above. In addition, figures (4-24), (4-25) showed the velocity rms estimation errors ($VQ_1(t)$, $VQ_2(t)$) calculated by equation (4-24) for TNF, EKF respectively.

Several comments can be drawn from Table 3 and Figures (4-18) through (4-25). First, Figures (4-18) - (4-21) showed that state estimates begin converging to their true values after own-ship executes a maneuver relative to target motion.

Figures (4-19), (4-21), (4-23), (4-25) showed the effect of increasing the measurements error to an rms value of 12^{o}. The EKF performance is improved while the TNF performance is degraded a little, but still the TNF performs significantly better than the EKF. The above conclusions are also demonstrated by Table 3.

Finally, it is generally noted that the covariance becomes smaller with decreasing range and grows as range increases. This was typical in all of the simulation cases, and pointed out the involvements of nonstationary processes. It is also, noteworthy to point out that when own-ship does not maneuver at all, both filters diverge (generating biased range estimates), but the EKF diverges faster than the TNF.

In closing, the following remarks are perhaps in order:

(i) In general, the TNF rms errors are much smaller than their corresponding EKF rms errors. This demonstrates that the TNF is more effective than the EKF.

(ii) Own-ship maneuvers relative to a constant-velocity target enhance convergence. Thus own-ship maneuver or target maneuver is essential to bearings only measurement analysis.

(iii) Computer utilization cost and storage requirements are almost equal for both algorithms. Thus computation efficiency is not a decisive factor in this application.

(iv) Finally, these results admittedly are not exhaustive. Thus, more simulation tests are needed especially for the case when the target is maneuvering. Also, comparison (via digital simulation) of the TNF with the MP (Modified Polar Coordinates) filter developed by Adiala [A3] warrant further consideration.

TABLE 4-1

Synopsis of the Percentage Accuracy of TNF over EKF

c_1	a	JJ_1%	JJ_2%
1.0	1	10.91	35.57
1.0	1.5	32.37	57.14
1.0	2	78.60	90.03
1.0	3	91.42	98.85

(σ_1= 2%, σ_2= 10%, $x_1(0) = 5\times10^3$m, $x_2(0) = 10^3$ m/sec)

TABLE 4-2

The Effect of Measurement Errors on thePercentage Accuracy of the TNF over the EKF

a	3	3	3
σ_1	2%	10%	20%
σ_2	10%	10%	10%
JJ_1%	91.42	11.0	-7.23
JJ_2%	98.03	17.47	-16.13

(c_1=1.0, $x_2(0)=10^3$ m/sec, $x_1(0)=5\times10^3$m)

TABLE 4-3

Synopsis of the Percentage Accuracy of TNF Over EKF as a Function of the Measurement rms Error

σ degree	2^0	3^0	6^0	12^0
JJ_1%	32.35	38.48	36.84	13.69
JJ_2	98.21	97.92	98.71	98.88

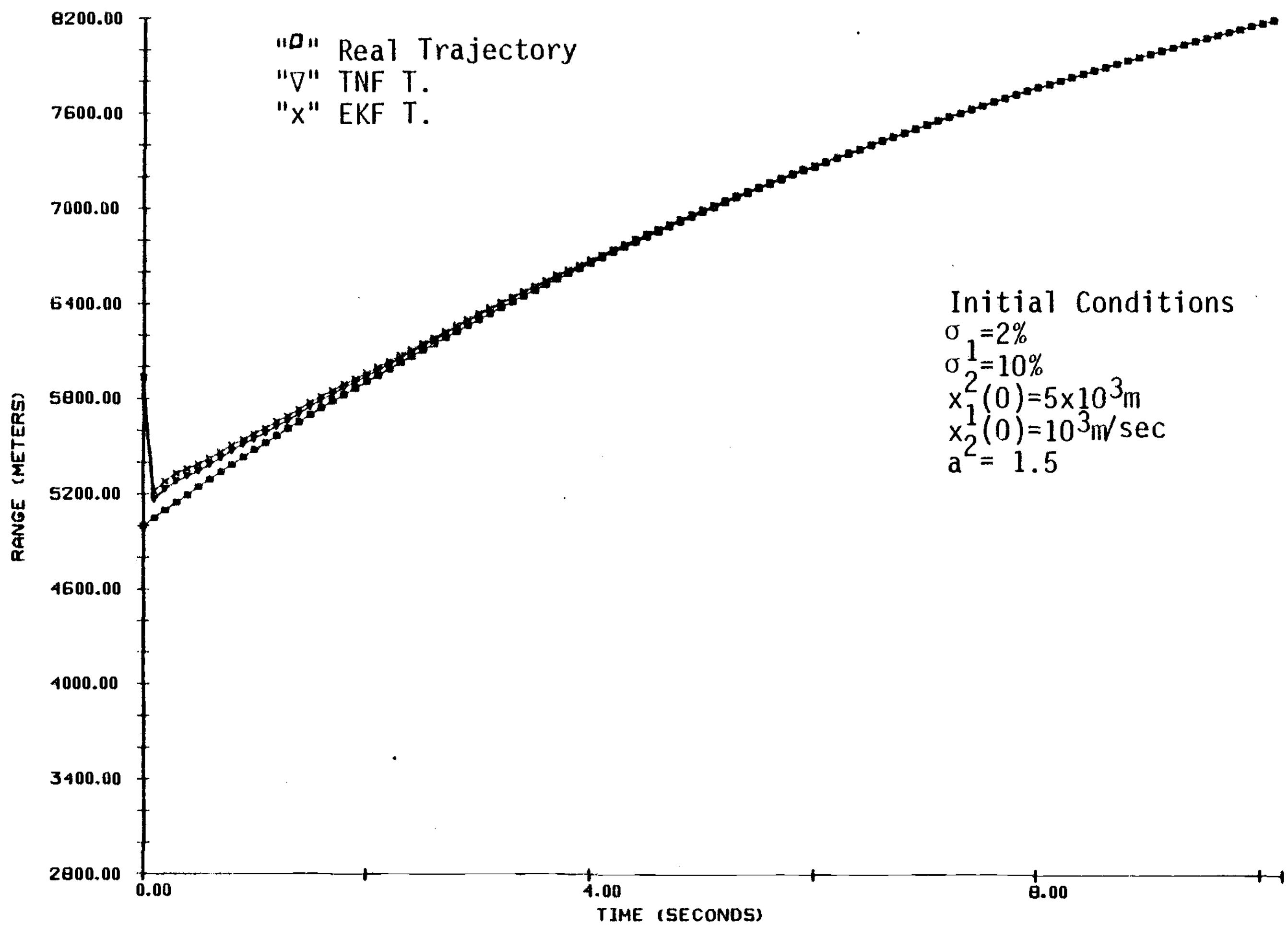


FIGURE 4-1 RANGE AND ITS ESTIMATES BY TNF,EKF

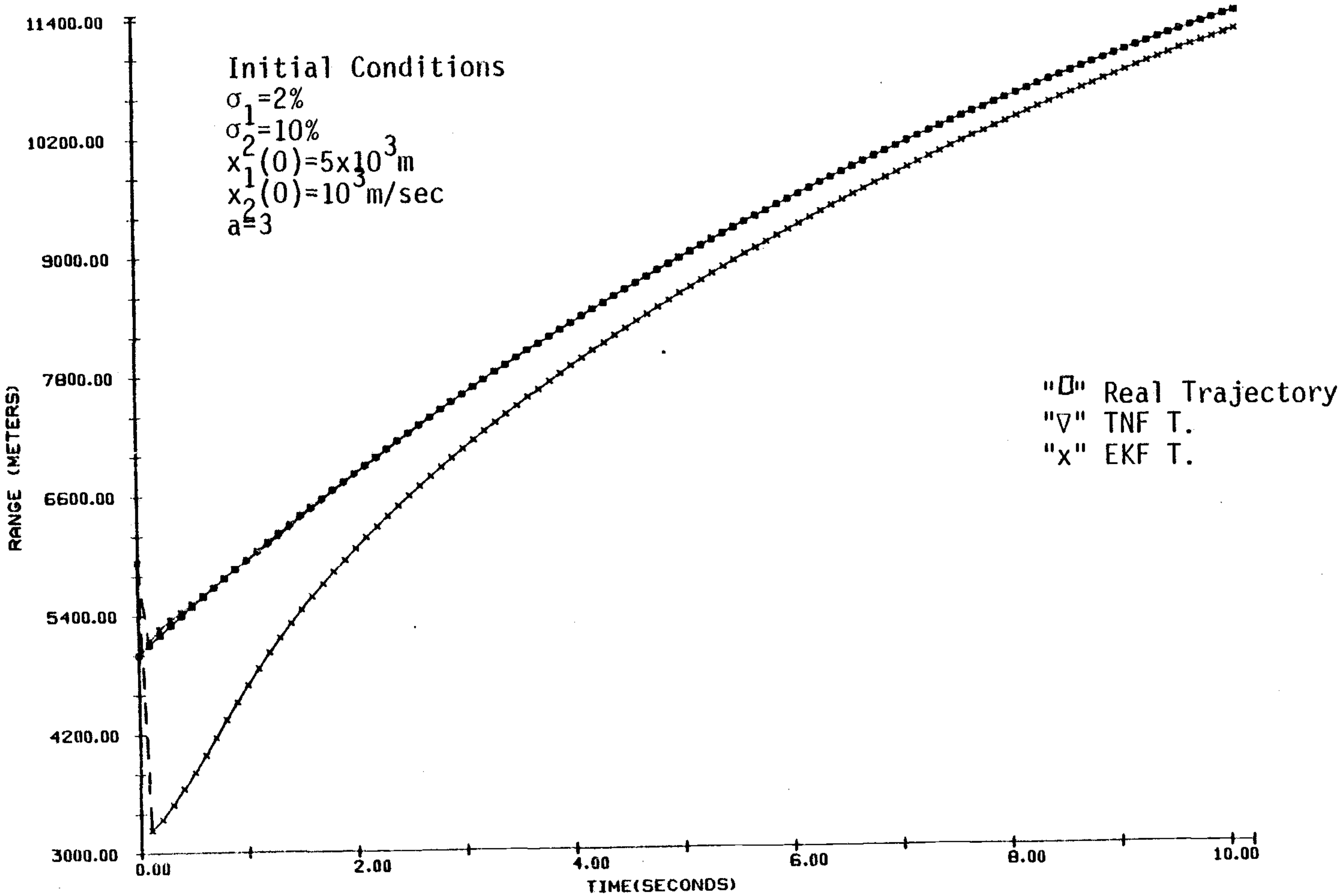


FIGURE 4-2. RANGE AND ITS ESTIMATES BY TNF,EKF

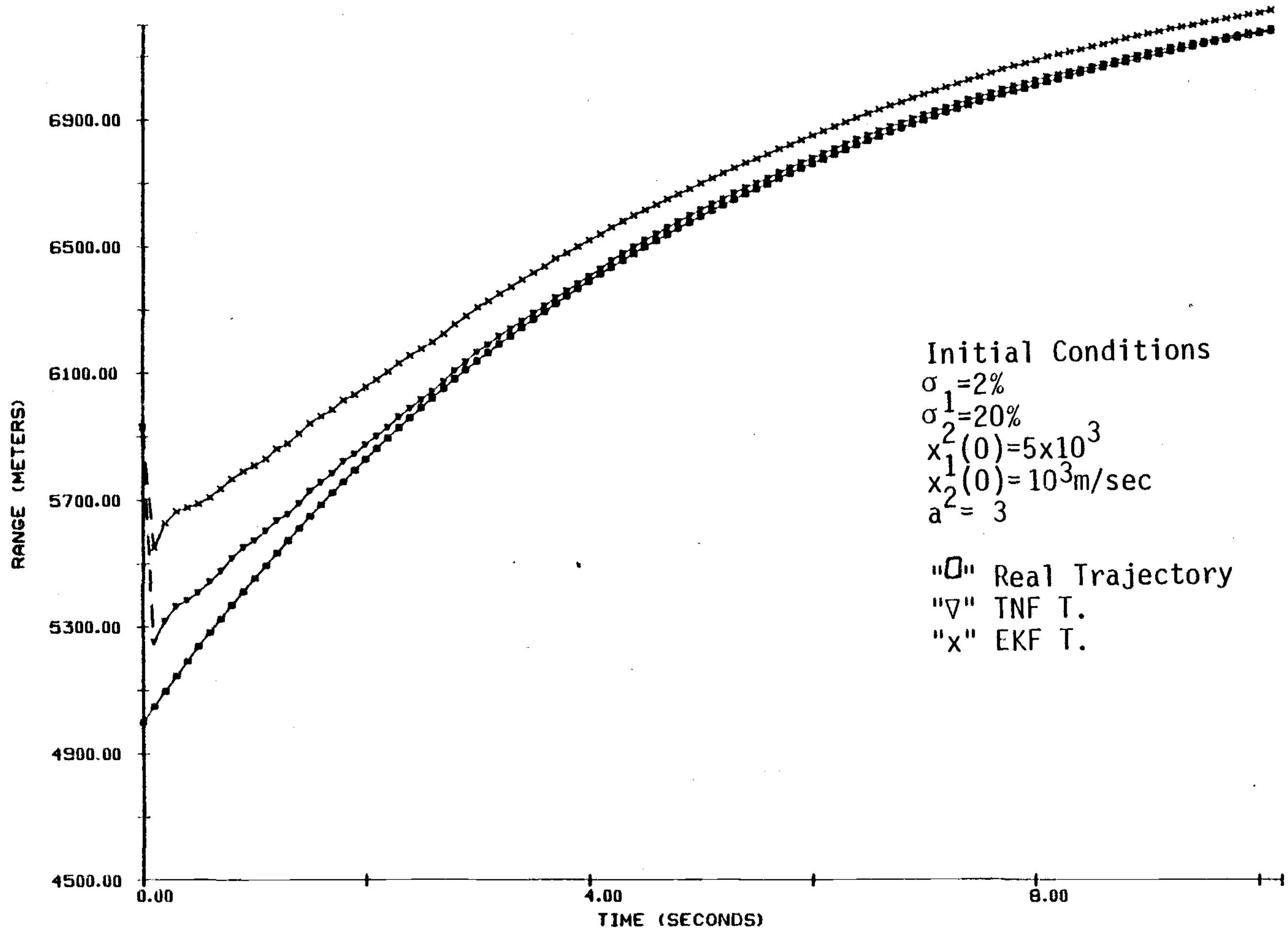


FIGURE 4-3 RANGE AND ITS ESTIMATES BY TNF,EKF

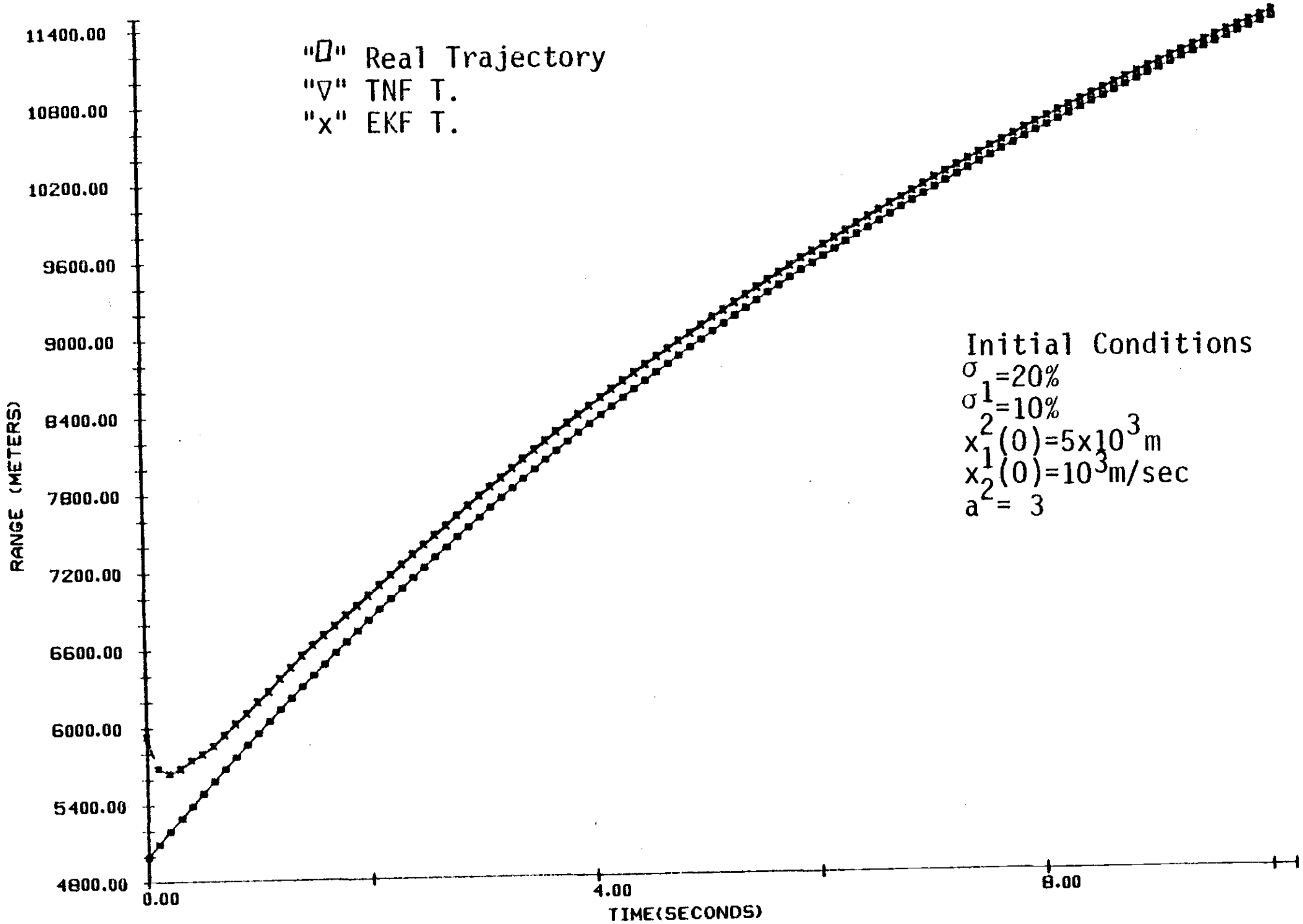


FIGURE 4-4 RANGE AND ITS ESTIMATES BY TNF,EKF

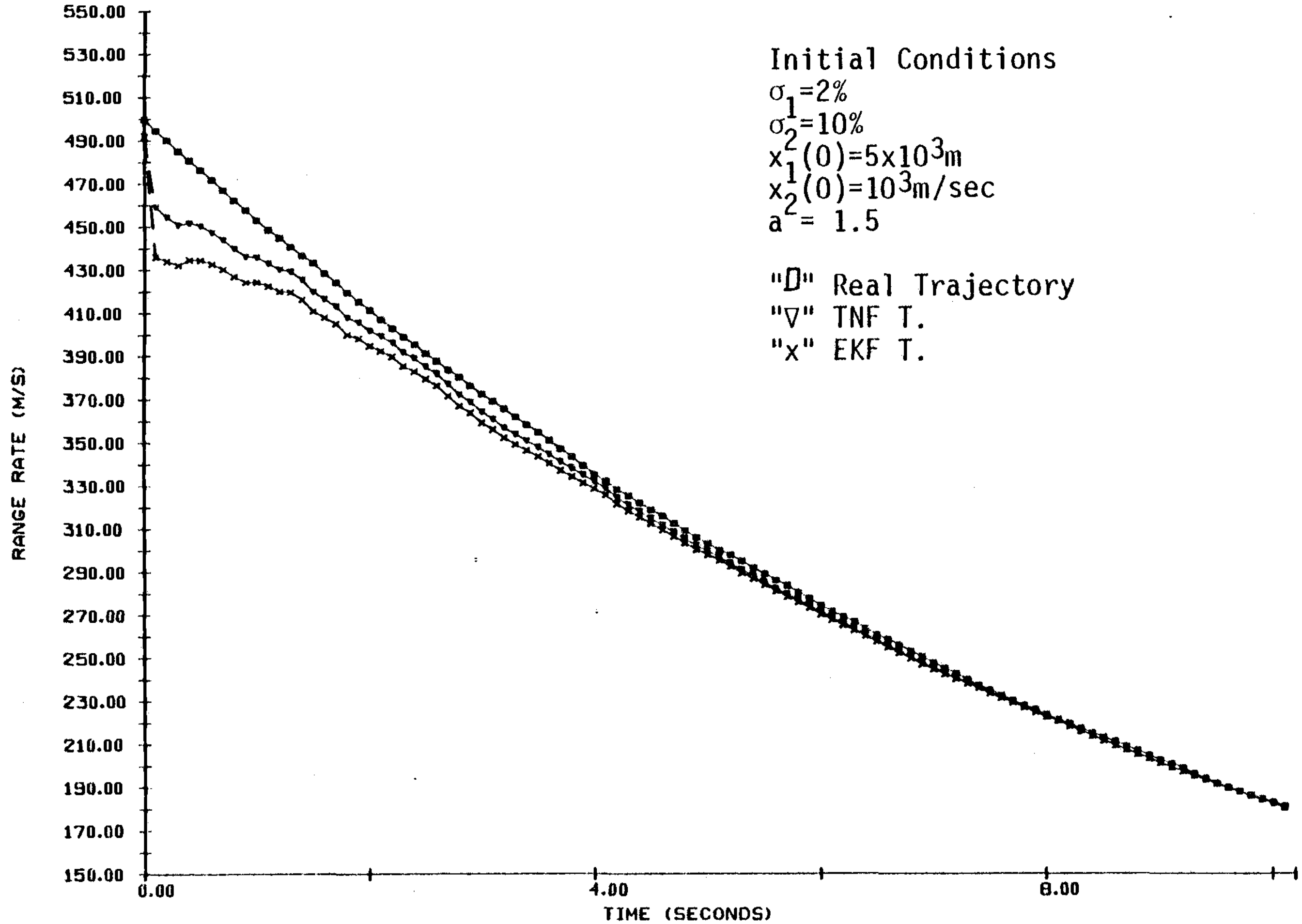


FIGURE 4-5 VELOCITY AND ITS ESTIMATES BY TNF,EKF

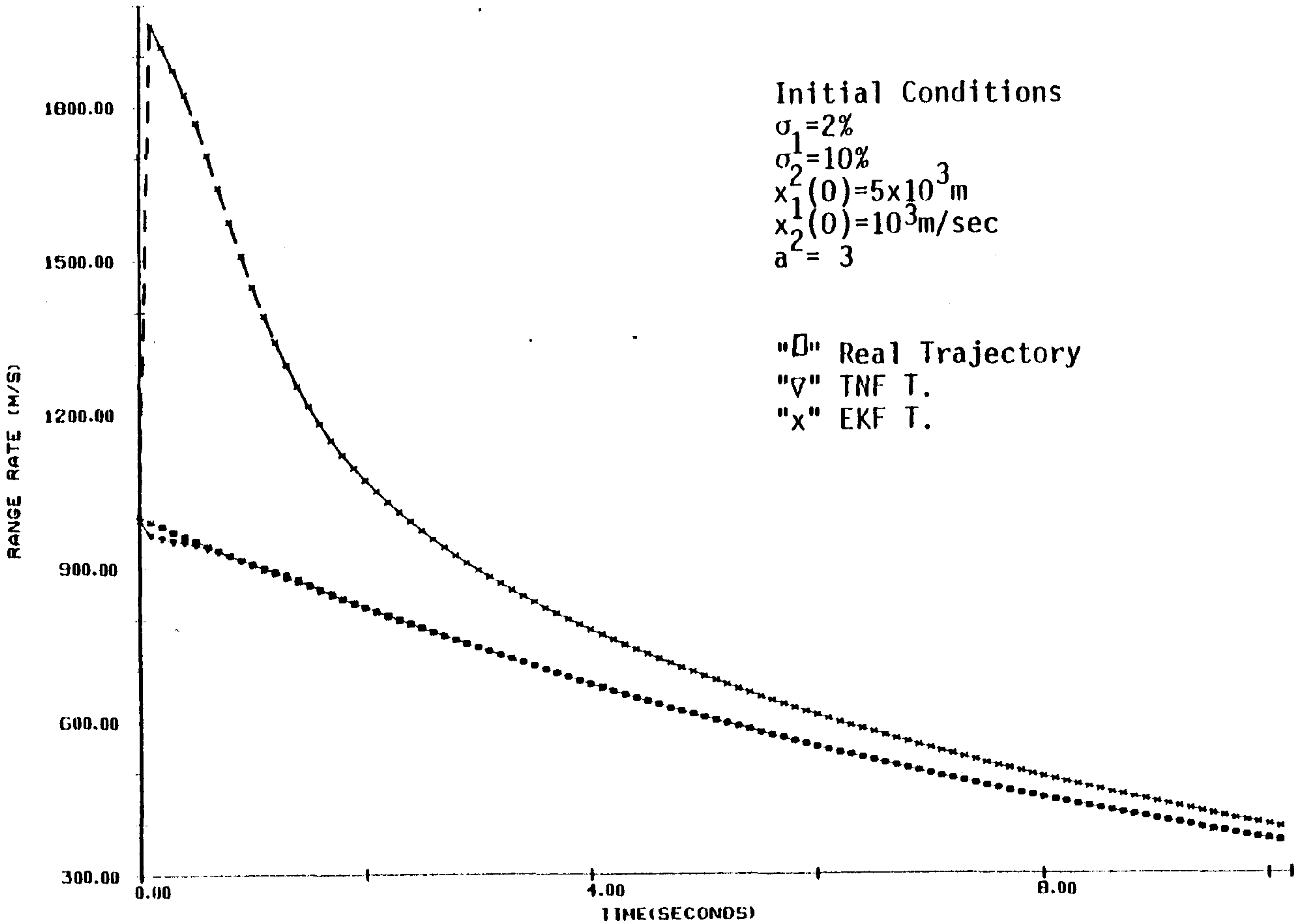


FIGURE 4-6 RANGE R.AND ITS ESTIMATES BY TNF,EKF

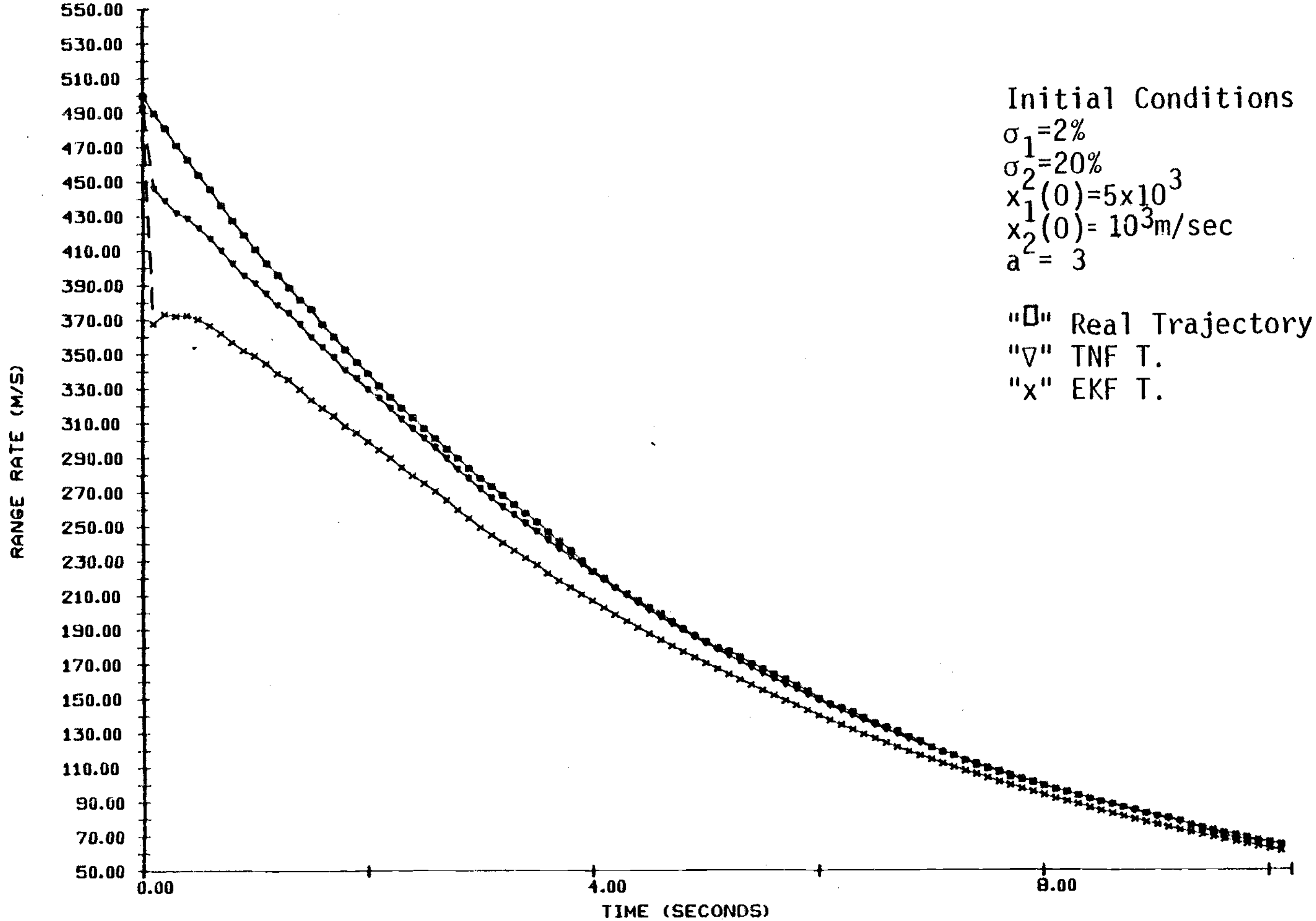


FIGURE 4-7 VELOCITY AND ITS ESTIMATES BY TNF,EKF

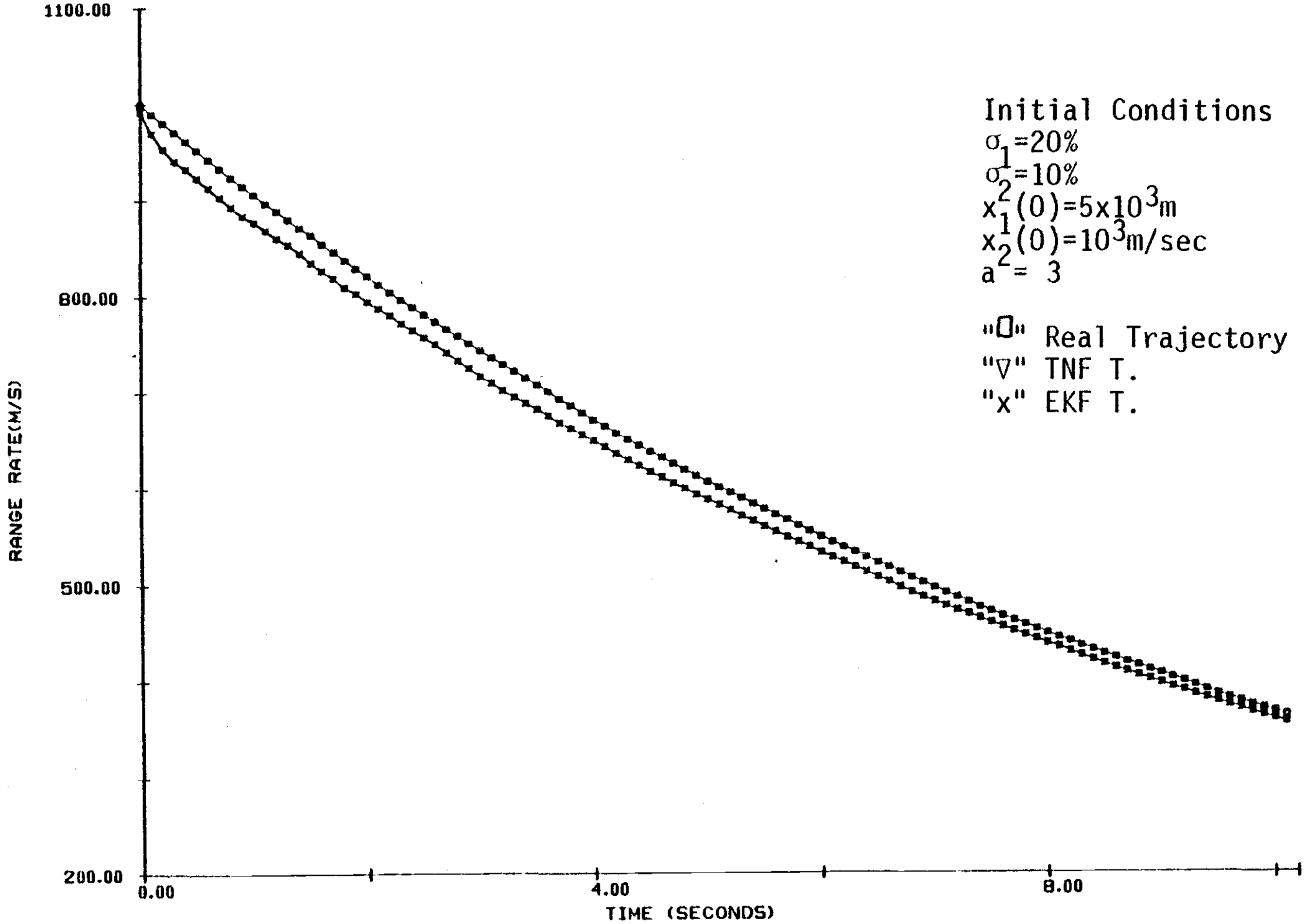


FIGURE 4-8 RANGE RATE AND ITS ESTIMATES BY TNF,EK

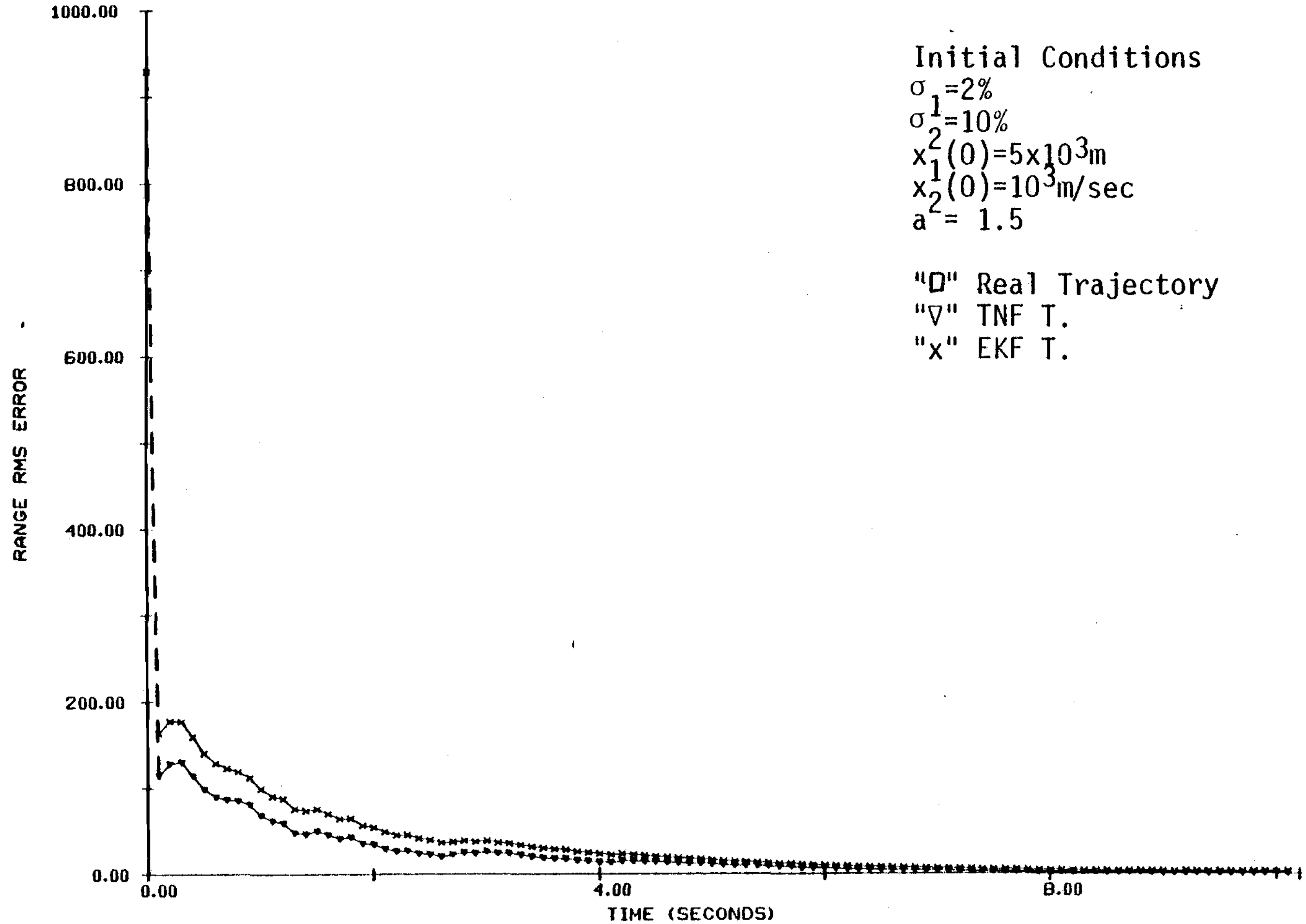


FIGURE 4-9 RANGE RMS ERRORS BY TNF,EKF

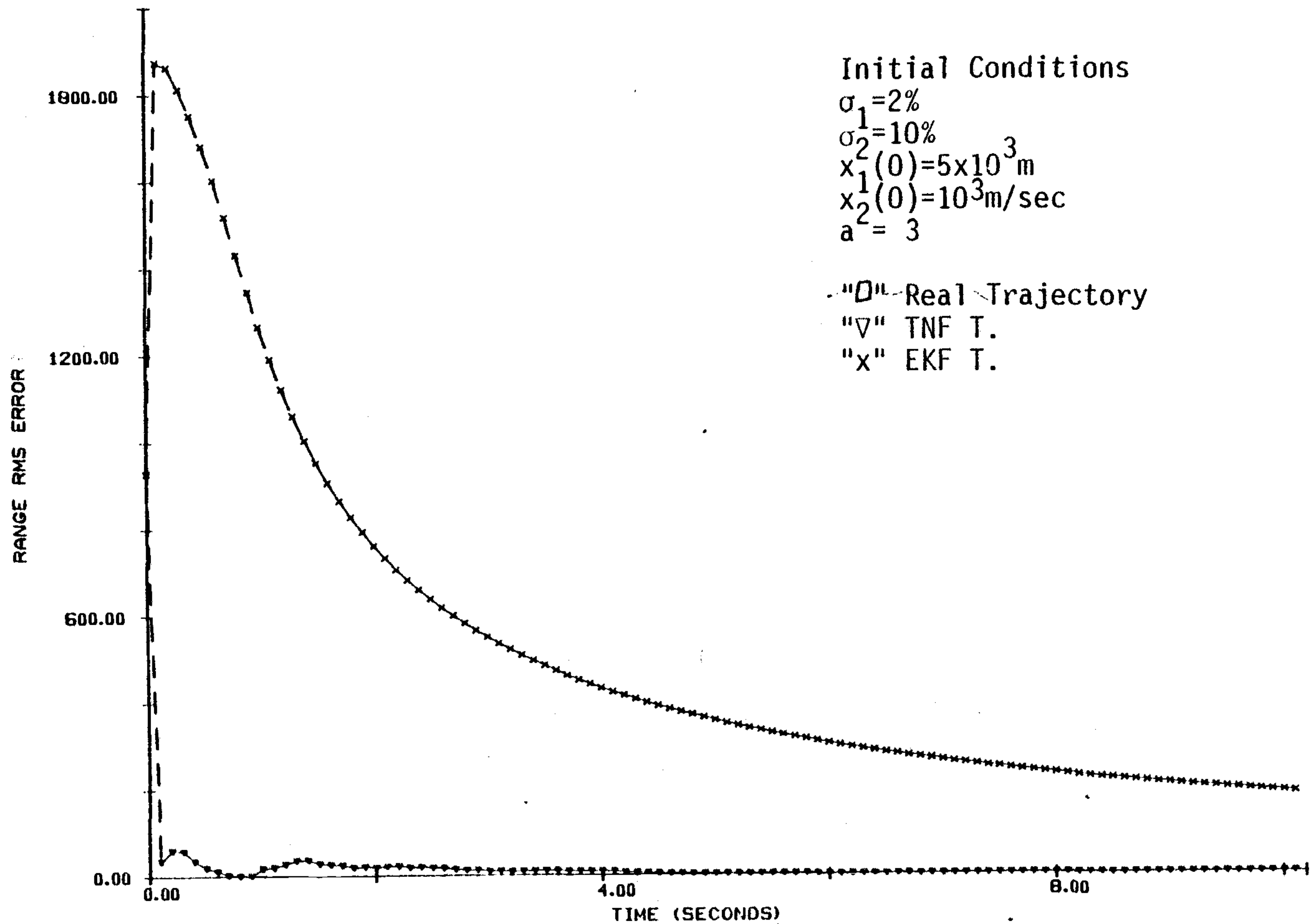


FIGURE 4-10 RANGE RMS ERRORS BY TNF,EKF

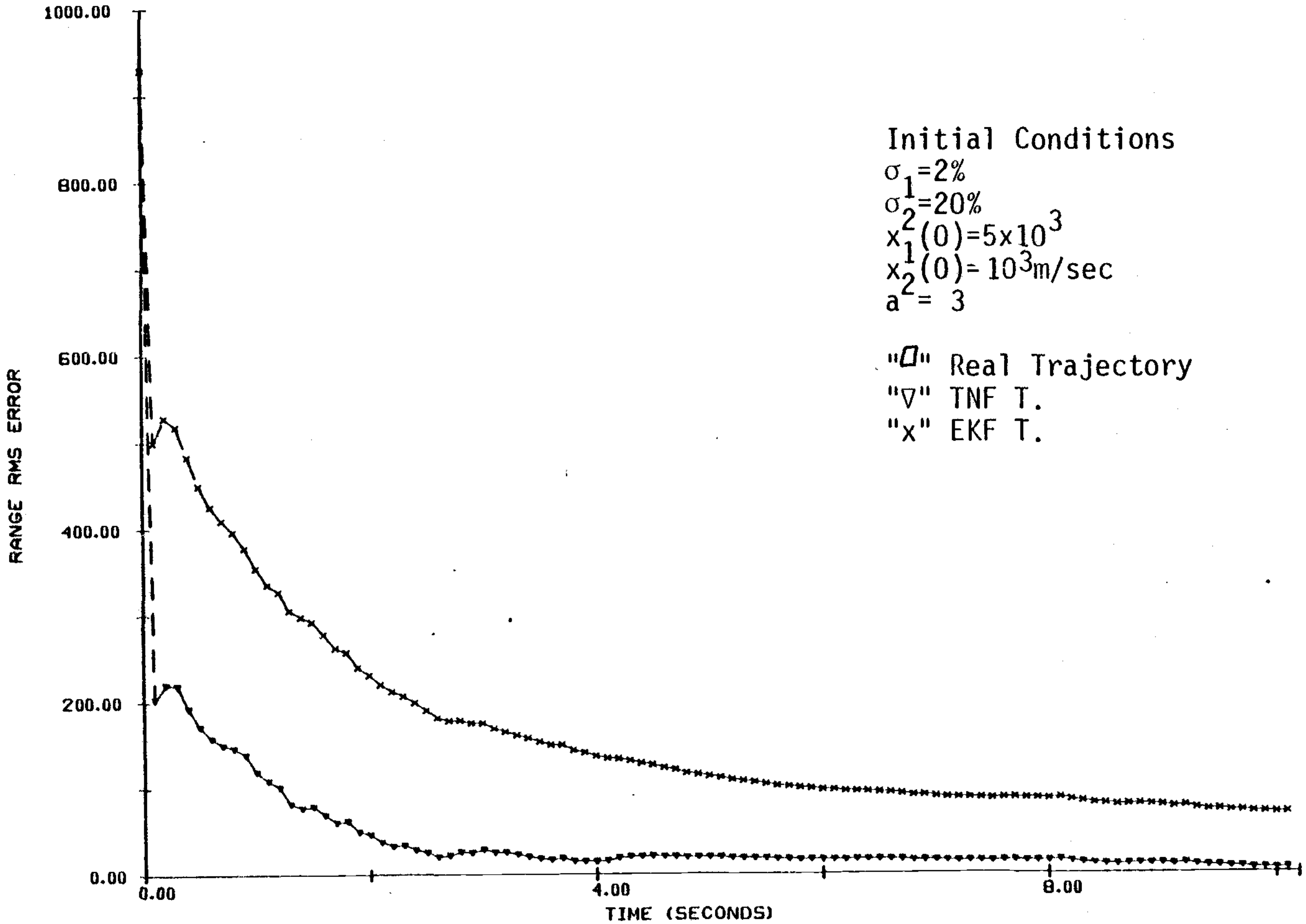


FIGURE 4-11 RAGE RMS ERRORS BY TNF,EKF

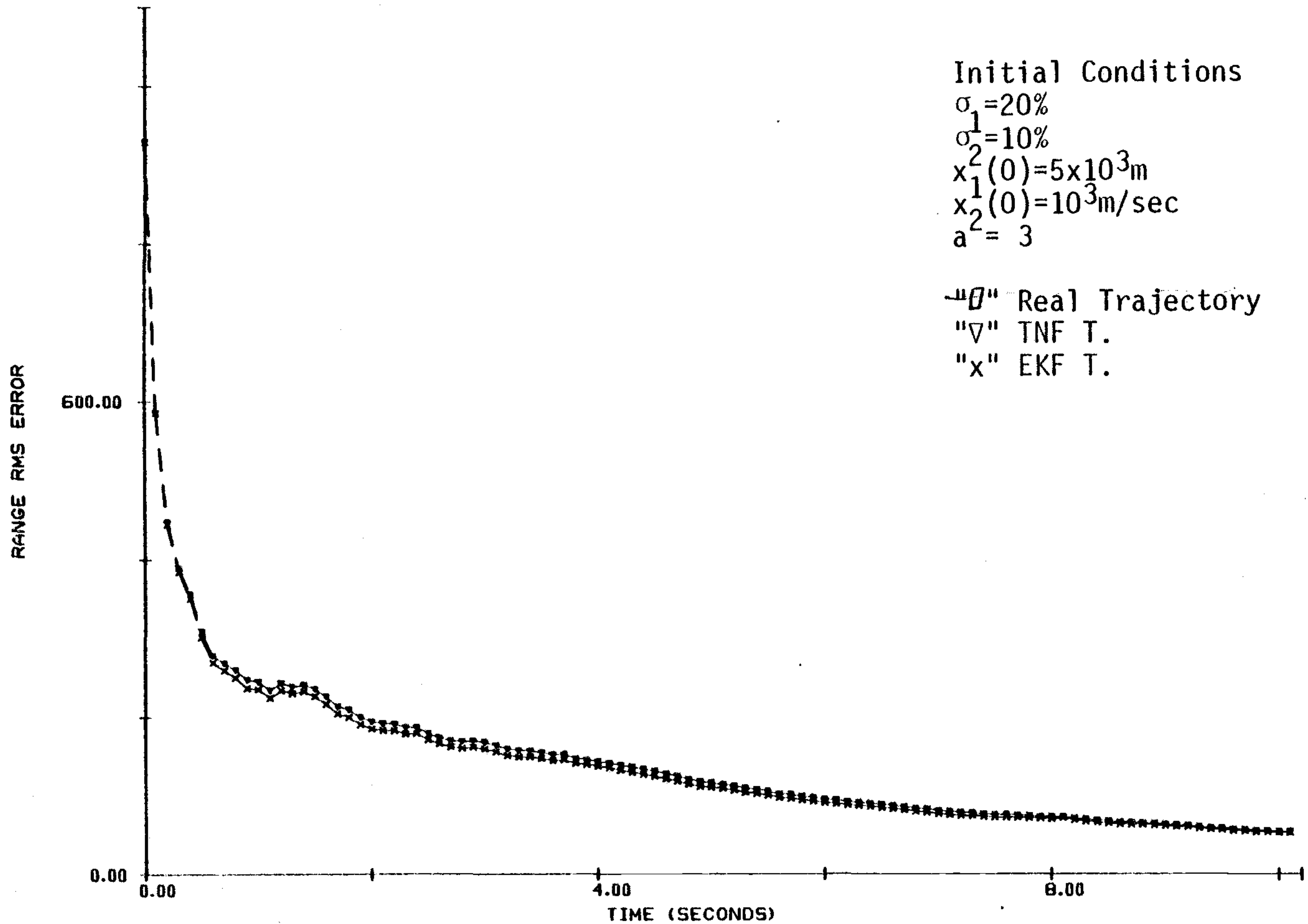


FIGURE 4-12 RANGE RMS ERRORS BY TNF,EKF

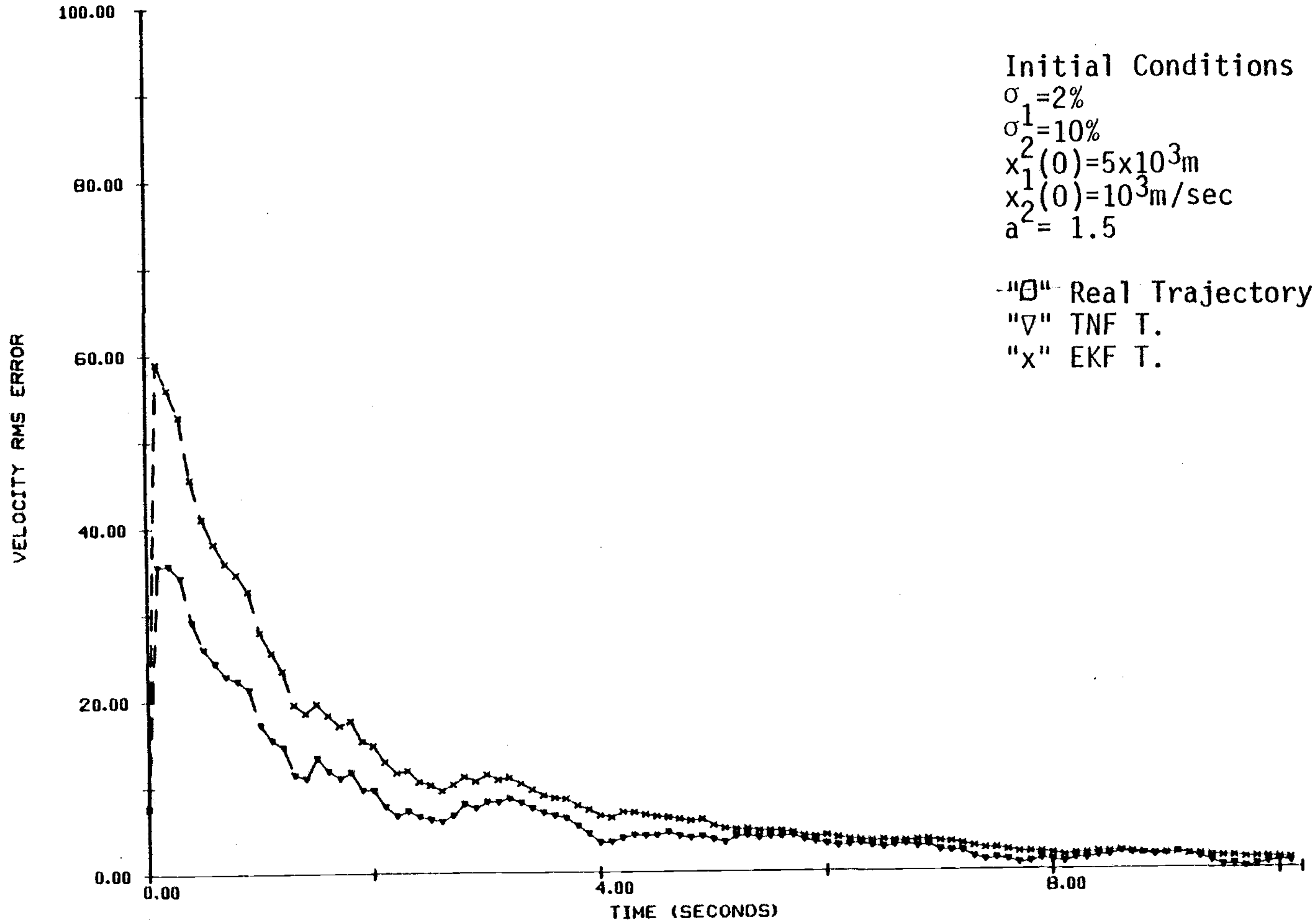


FIGURE 4-13 VELOCITY RMS ERRORS BY TNF,EKF

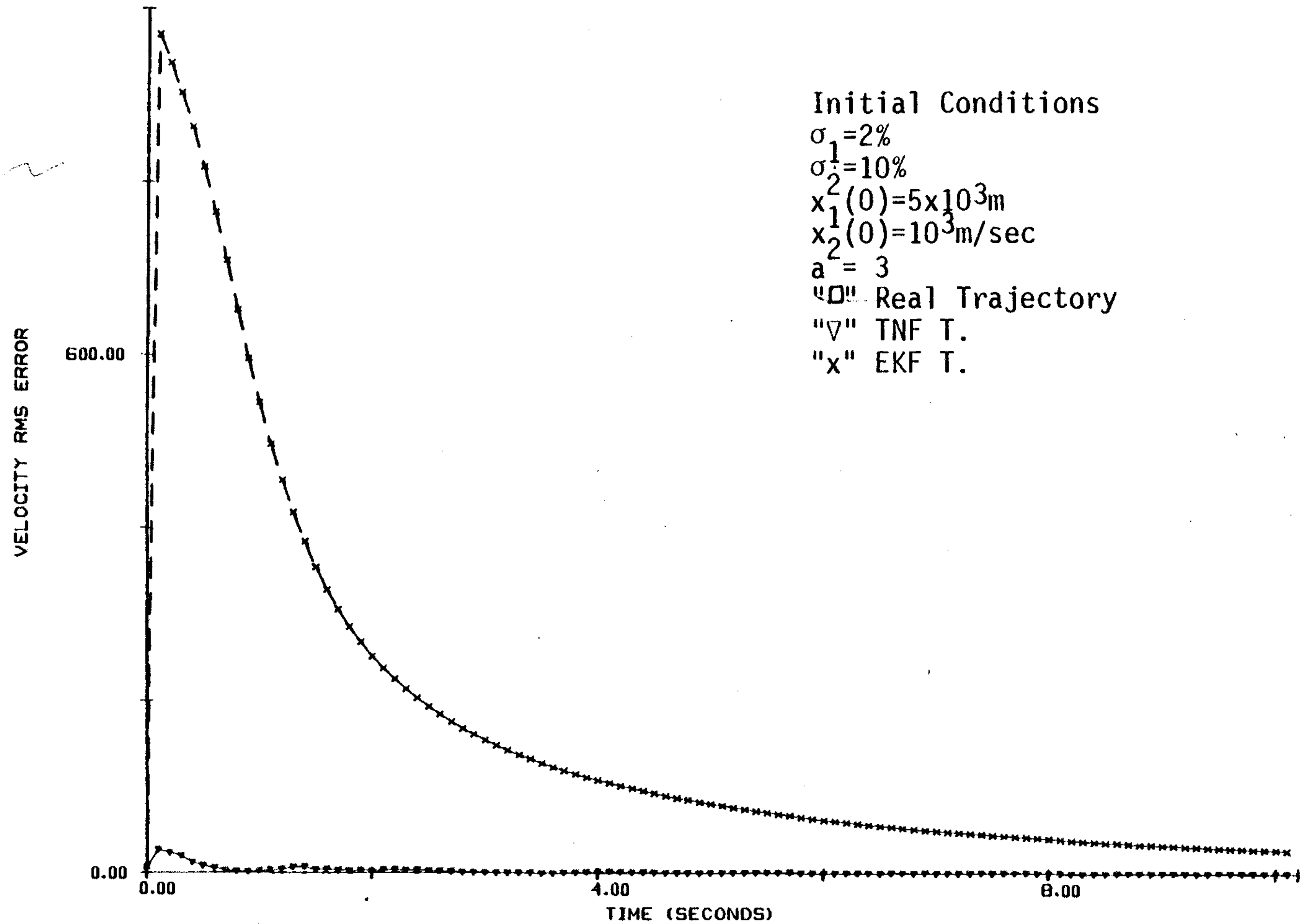


FIGURE 4-14 VELOCITY RMS ERRORS BY TNF,EKF

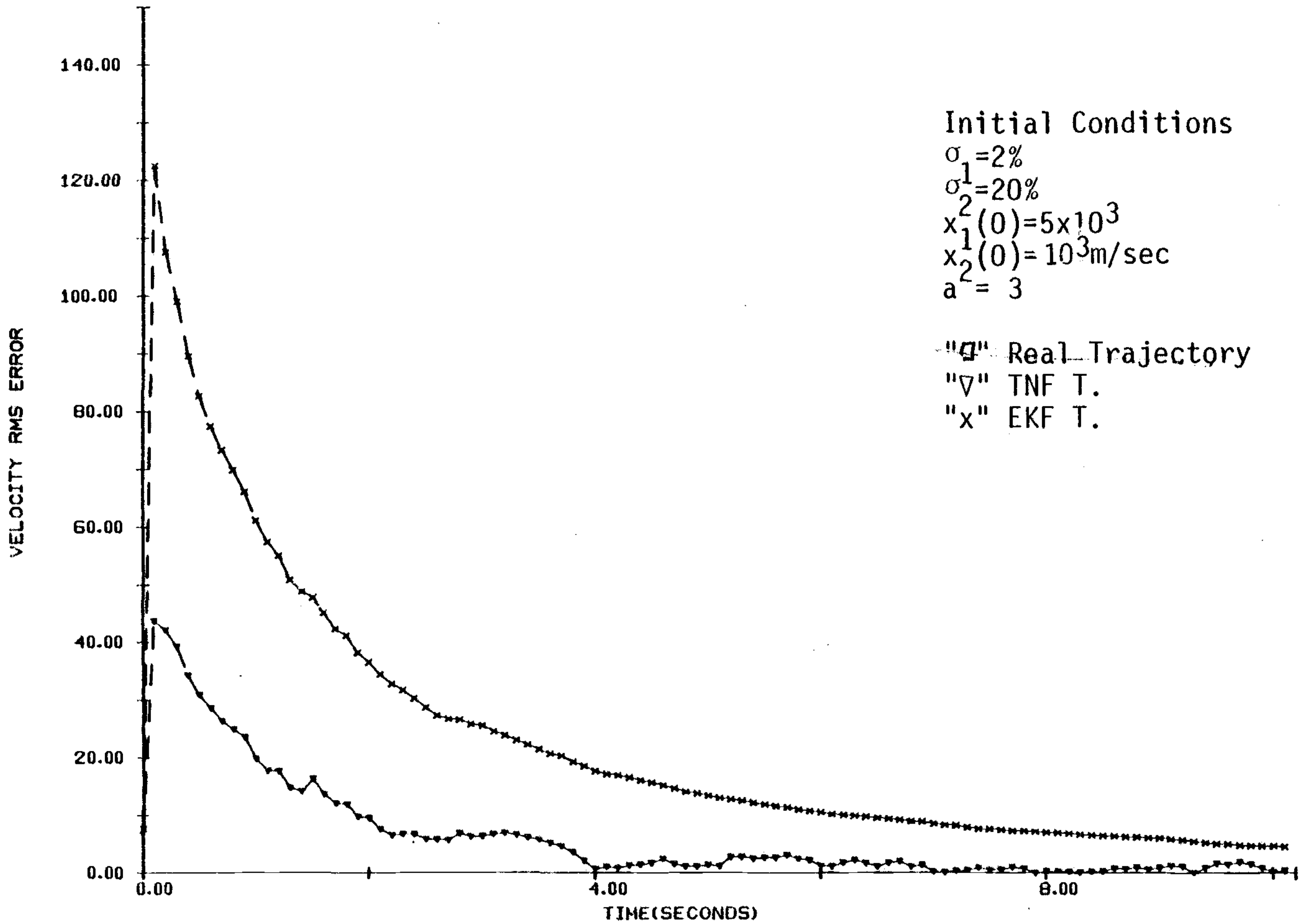


FIGURE 4-15 VELOCITY RMS ERRORS BY TNF,EKF

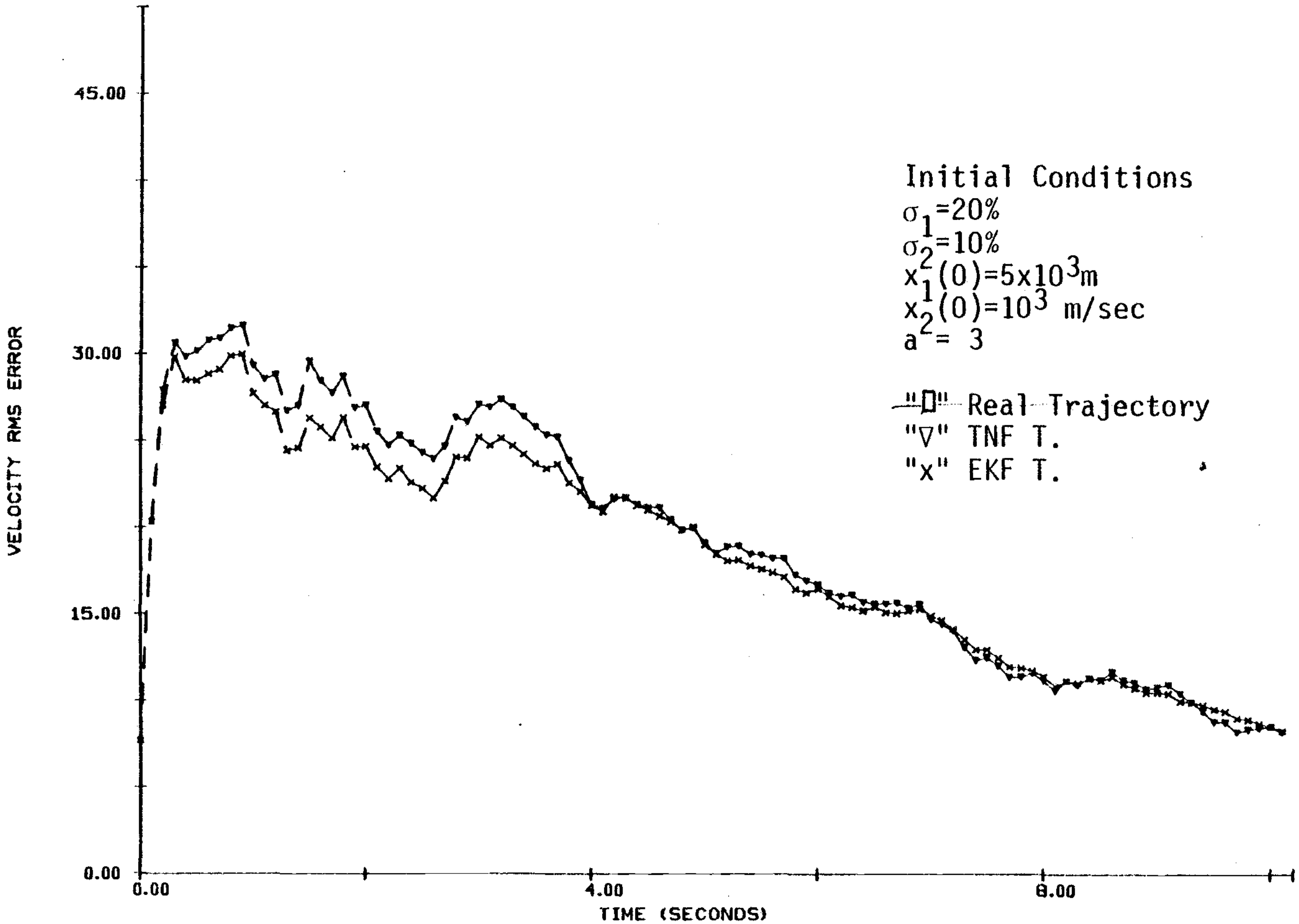


FIGURE 4-16 VELOCITY RMS ERRORS BY TNF,EKF

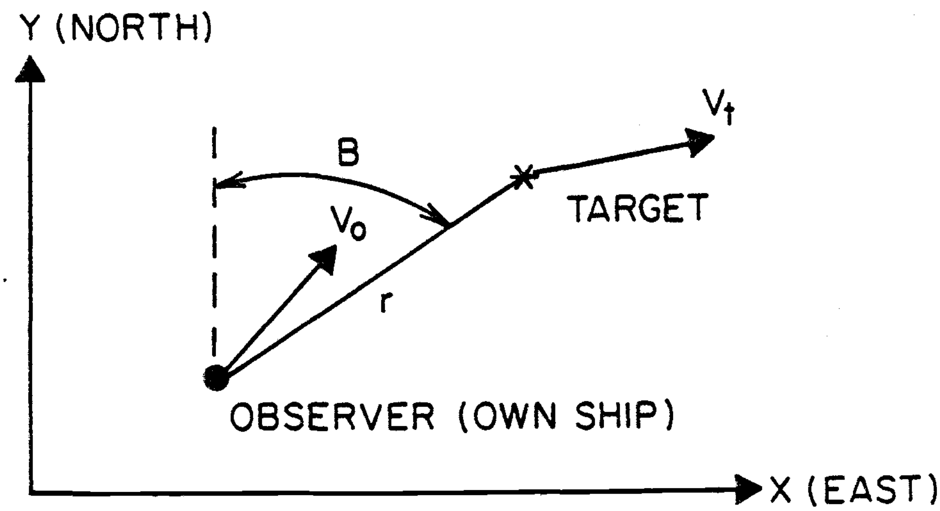


Figure 4-17

Geometric Configuration for Bearings-only Motion Tracking

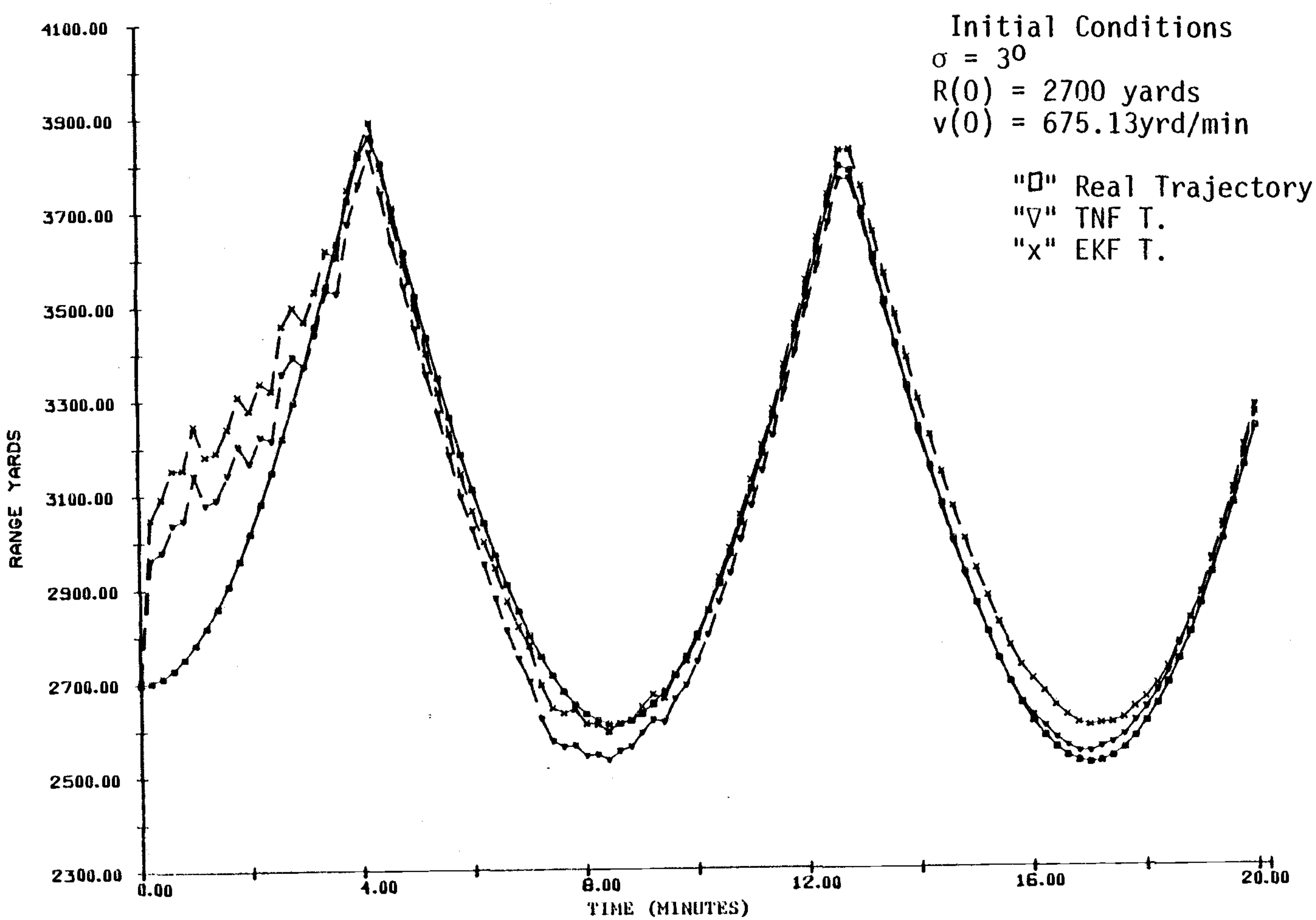


FIGURE 4-18 RANGE AND ITS ESTIMATES BY TNF,EKF

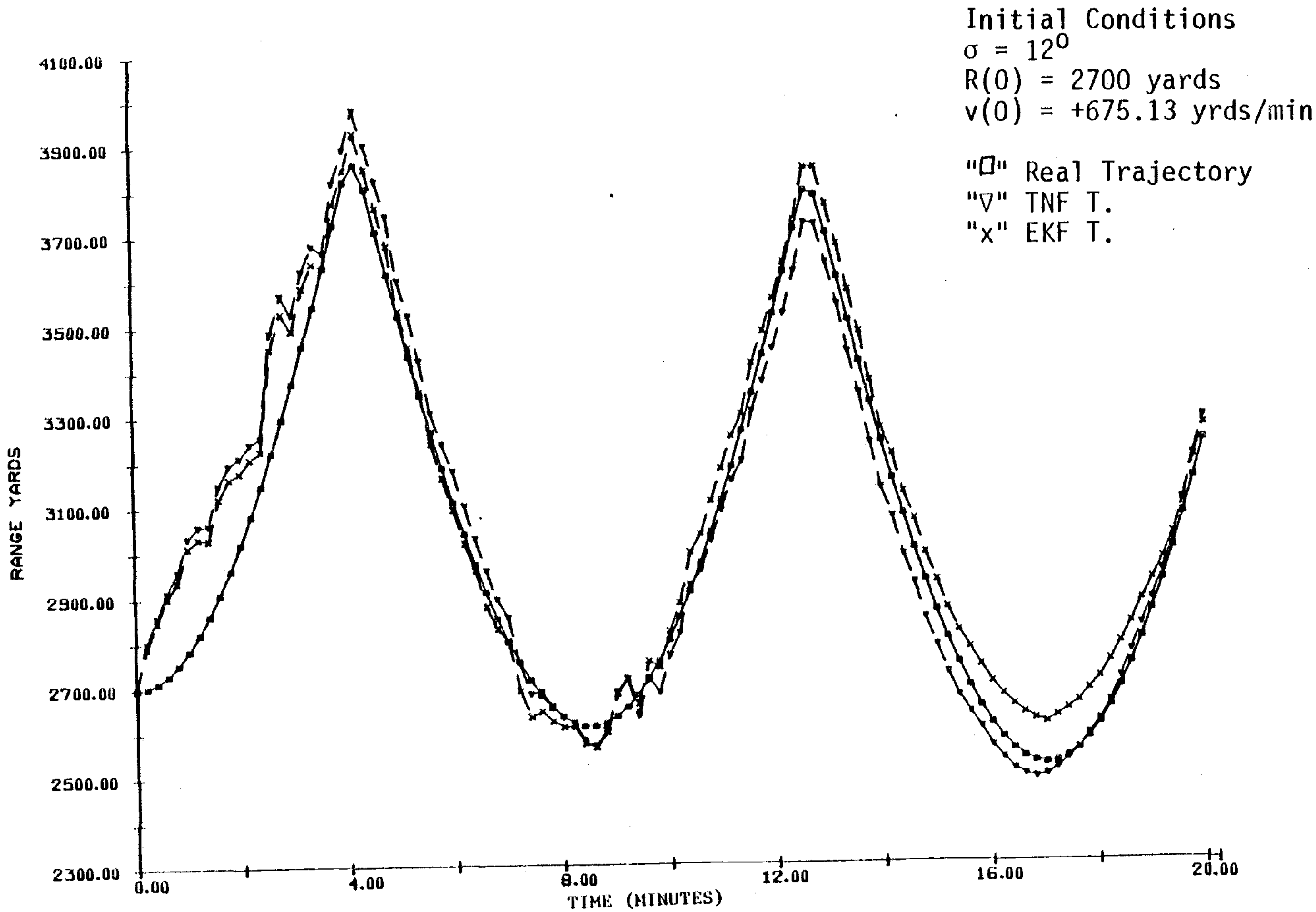


FIGURE 4-19 RANGE AND ITS ESTIMATES BY TNF,EKF

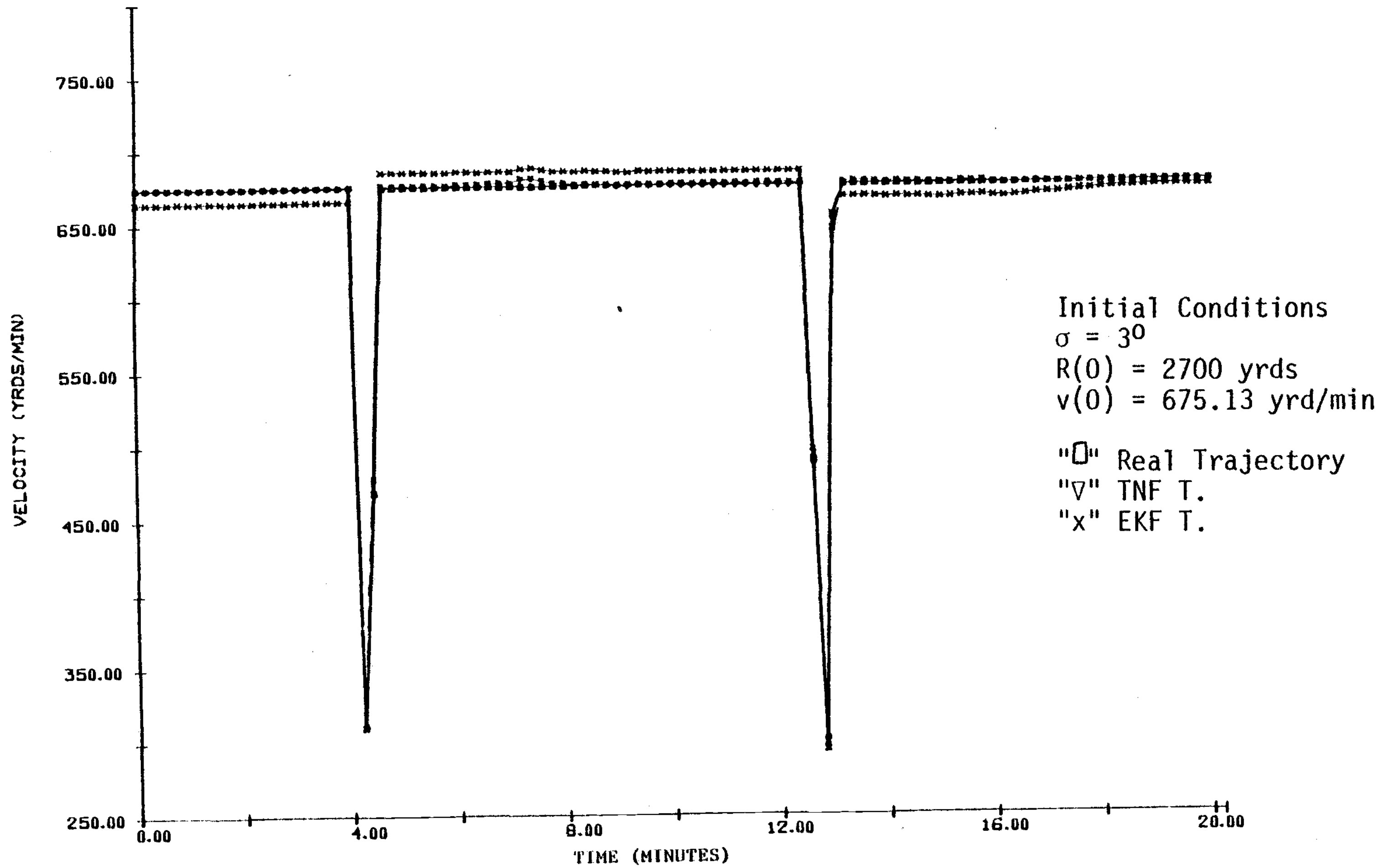


FIGURE 4-20 VELOCITY AND ITS ESTIMATES BY TNF,EKF

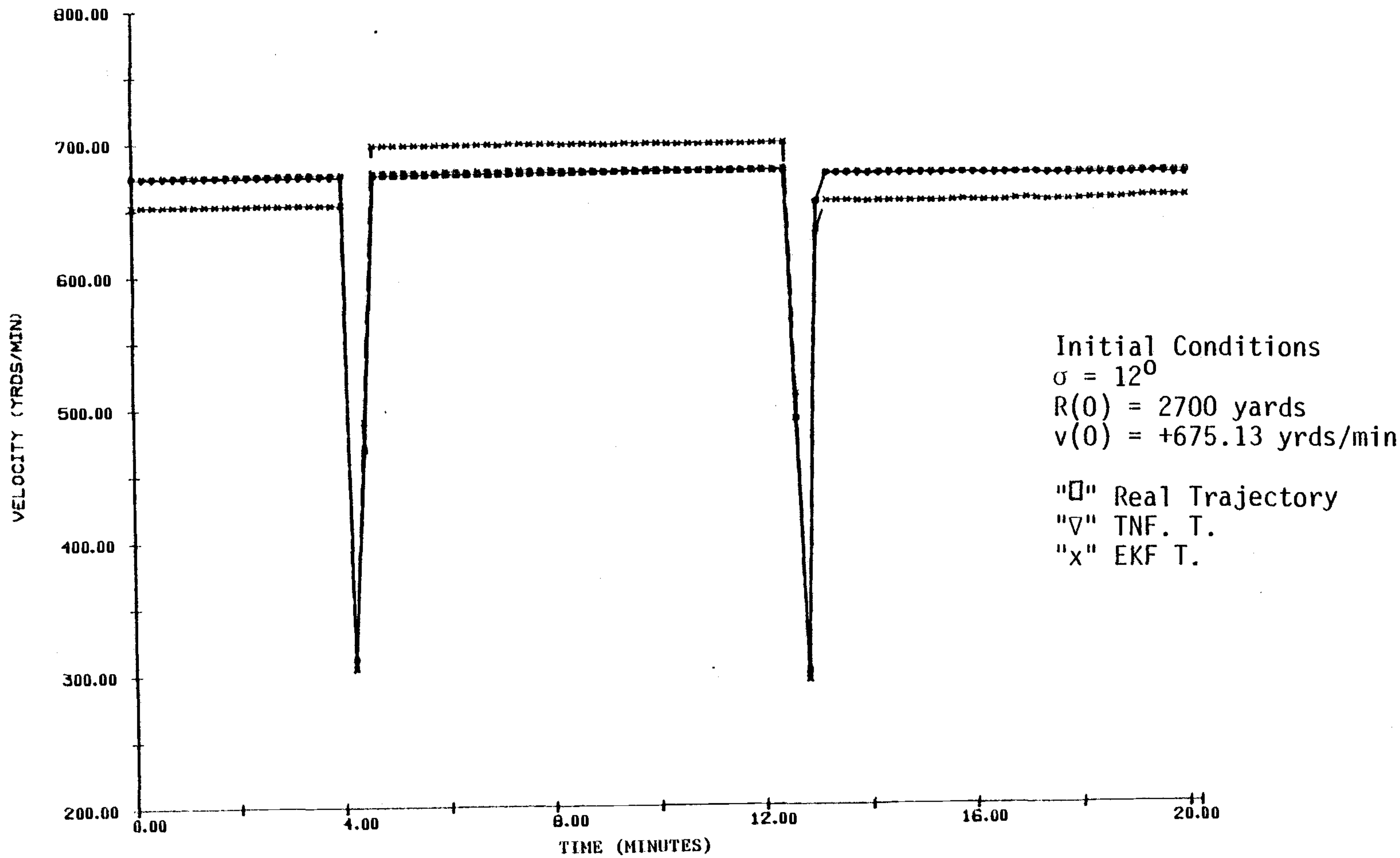


FIGURE 4-21 VELOCITY AND ITS ESTIMATES BY TNF,EK

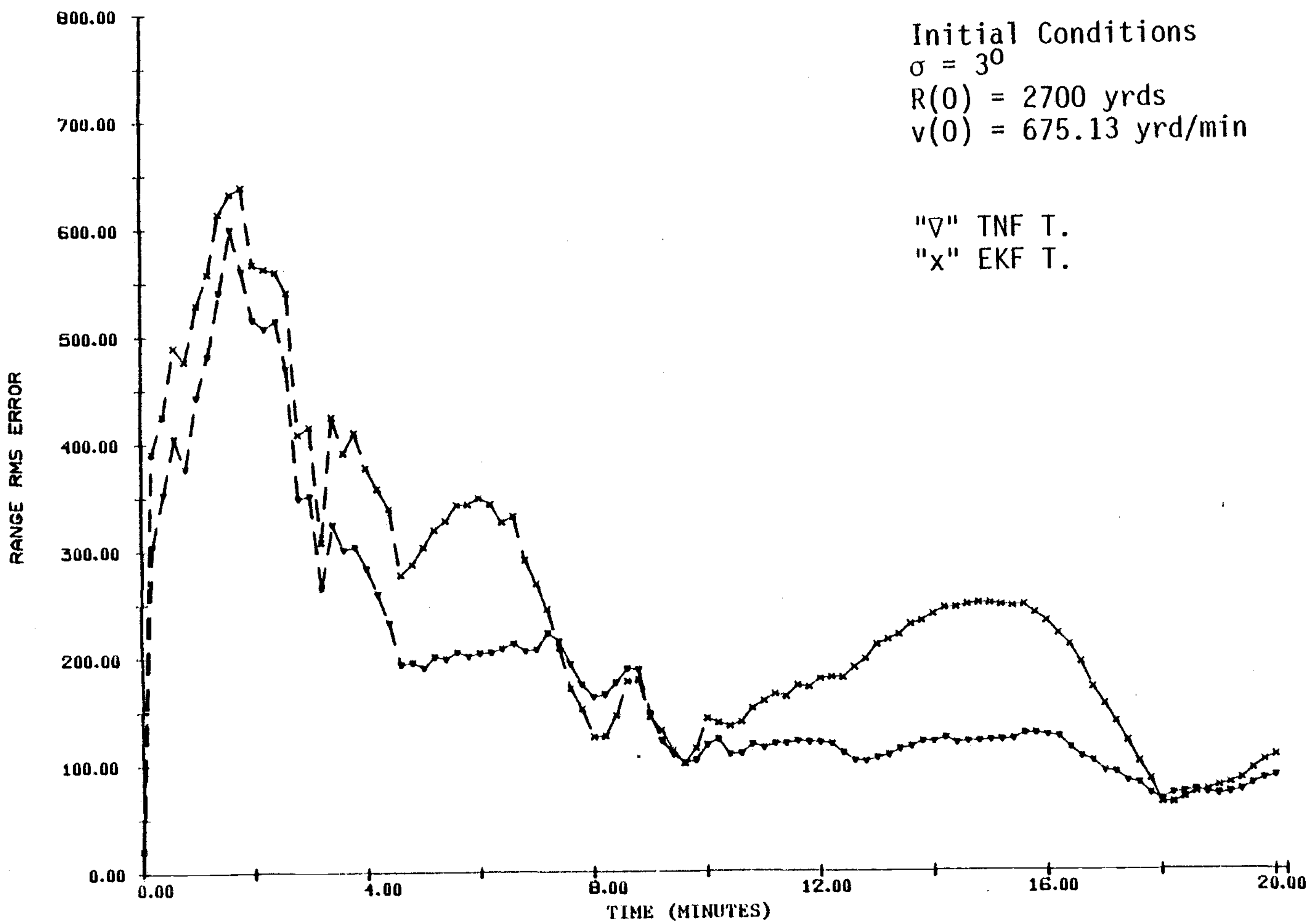


FIGURE 4-22 RANGE RMS ERRORS BY TNF,EKF

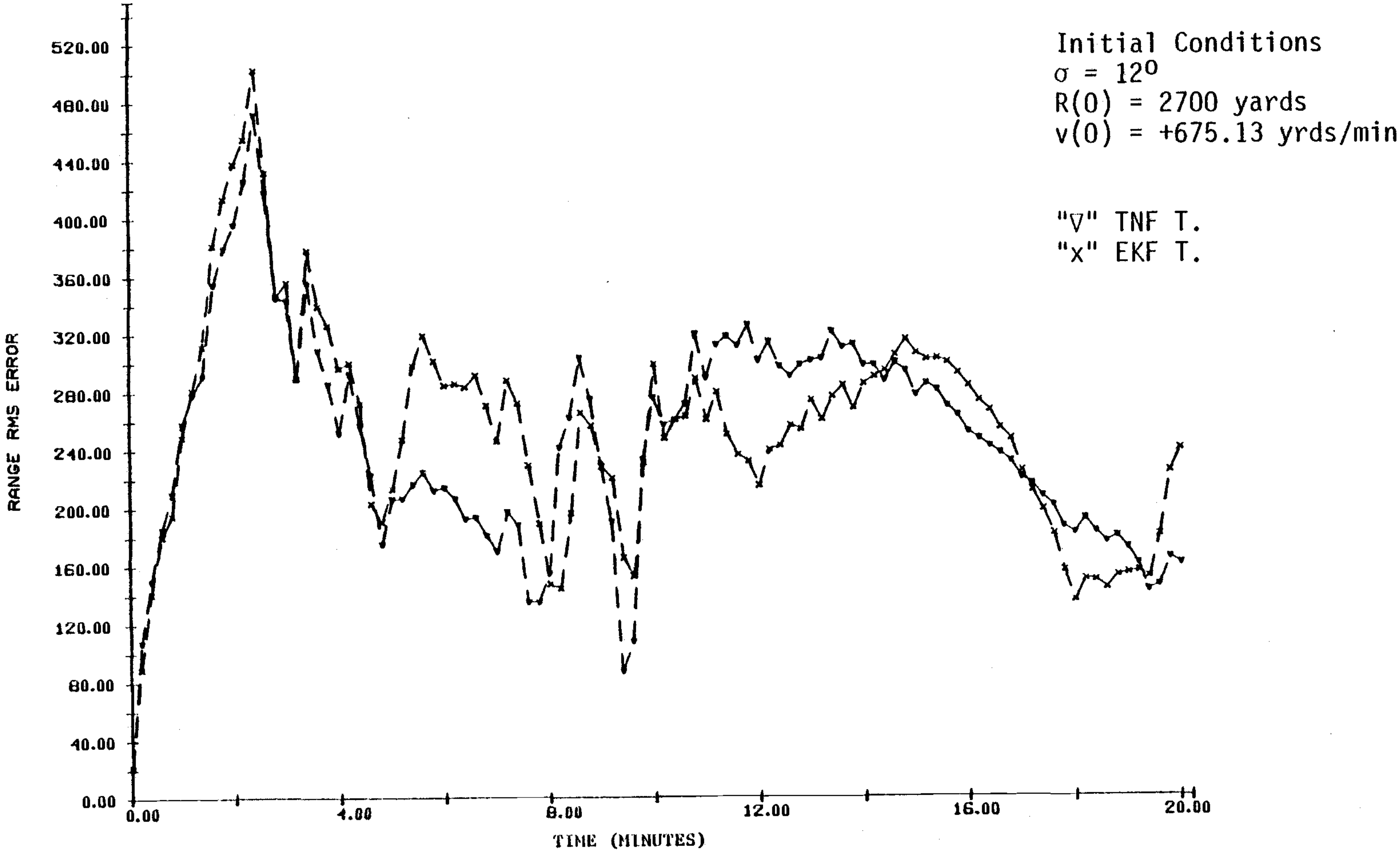


FIGURE 4-23 RANGE RMS ERRORS BY TNF,EKF

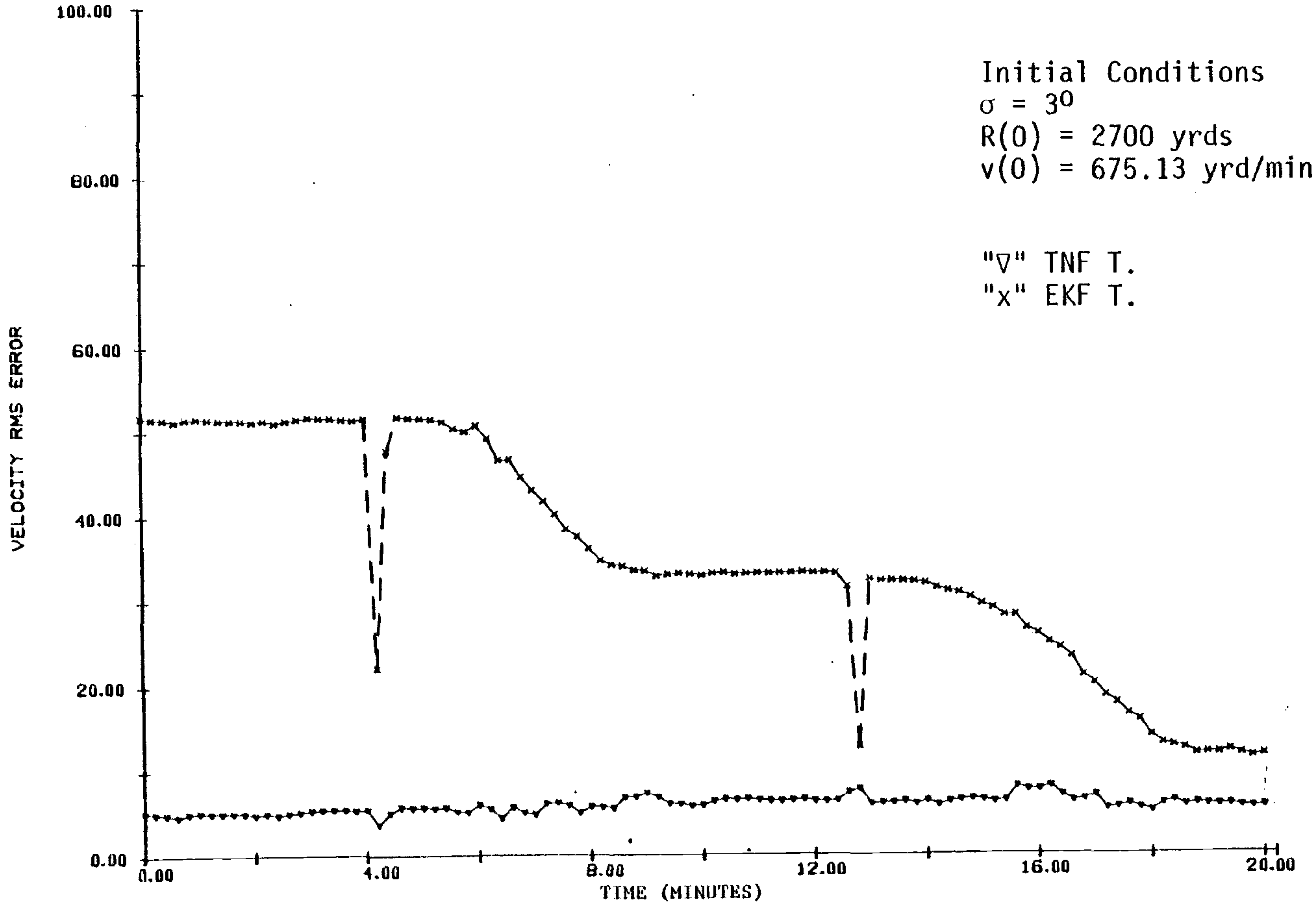


FIGURE 4-24 VELOCITY RMS ERRORS BY TNF,EKF

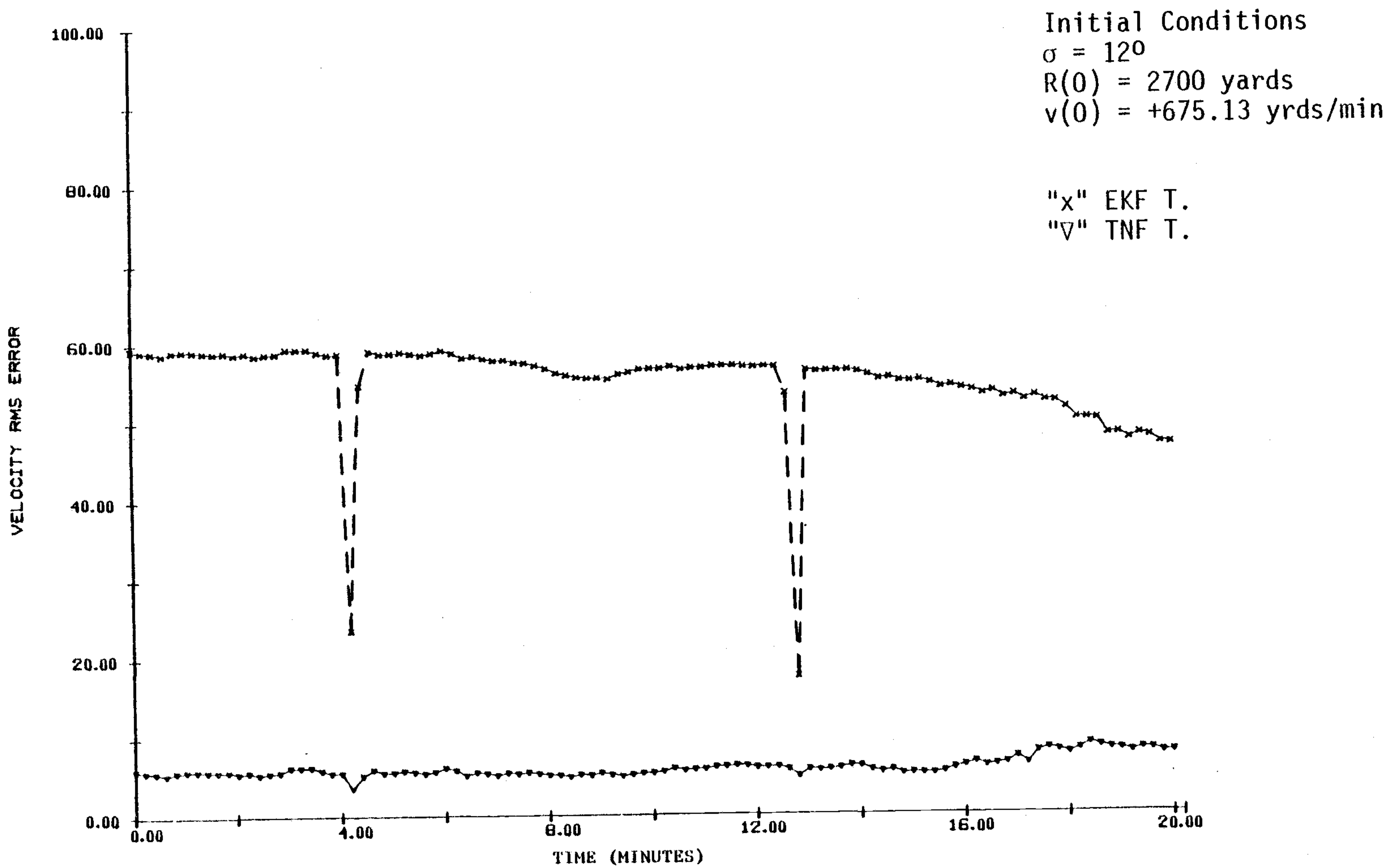


FIGURE 4-25 VELOCITY RMS ERRORS BY TNF,EKF

Chapter 5

5. SUMMARY AND CONCLUSIONS

In this research an original nonlinear filter approximation is developed for a class of nonlinear systems. The need for such a filter which does not suffer from the shortcomings of most of the linearization techniques, such as model smoothness, is encountered frequently in many industrial, sonar, economic, and image-processing applications. Obviously, the design of a filter which has improved performance, without stringent requirements, and which is employed with comparable implementation cost to the traditional techniques, would be a very significant achievement. This has been the major design goal of the work presented herein.

5-1 Significant Features of the New Filter

The effort was devoted to the development of a new finite-dimensional filtering approximation to the typical infinite-dimensional-nonlinear filtering problem. As a result of this effort, new features were developed, and a modest contribution to nonlinear-filtering approximation theory was achieved. Among these are the following:

1) Weaker assumptions are required to derive the filter even as the solution is assumed to be a weak solution.
2) Impeding generation of a "close", (in m.s.e), bilinear, feedback-control-model approximation to the original nonlinear system.

This has an impact on nonlinear approximation theory, since it provides formal approximation procedures with flexibility embedded. That is due to the fact that the appproximating parameters are functionals of the feedback control u_t which provides some control and insight into the structure of the approximating model.

3) The feedback coupling in the filter-covariance equations enhances the stability of the filter. In all of the linearization techniques only forward coupling occurs in the filter equations and that in essence helps to destablize the filter due to error accummulation.

4) An important feature of the new filter is the "global" property which allows the filter to be independent of the local-time discretization, and thus decreases the noise aggravation.

5) A new "decomposition scheme" is developed for a certain class of large-scale nonlinear systems. This results in great mathematical simplification in the development of global nonlinear multi-dimensional filter.

6) This filter provides an alternative formal approach to certain traditional techniques, (such as MSOF [J1], and "the Smoothness in Probability Filter [K4]), for certain nonlinear systems even though the approaching method is quite different.

5-2 Future Area of Research

The following five main areas warrant further study:

1) The development of an equivalent filtering algorithm for the case when the observation process is discrete. This is specifically important in sonar applications since observations may be received as batches at random times. The following cases are of most interest:

   a) the observation is discrete but random;

   b) both the observation and the system processes are discrete and random.

2) Investigation of a general "Decomposition Scheme" to extend the applicability of the new filter to a broad class of nonlinear systems. The use of a hierarchical optimality scheme is promising.

3) Computer utilization is an issue that should be more fully explored. Thus, development of a standard software package for the TNF algorithm would make the filter quite attractive to practical applications. Furthermore, incorporation of the microprocessor would increase the economical feasibility of the filter.

4) The need for further testing and comparison with other existing techniques to establish filter superiority in most cases, and to enhance its reliability.

5) Investigation of the use of "adaptive" techniques to surmount the difficulty of tracking successfully an evasive maneuvering target, which is a typical and important subject in sonar applications. The following techniques are promising:
   a) use of an additive control that can be estimated separately using the separation principle for stochastic conditionally Gaussian systems [K1];
   b) use a feedback scheme to correct for the target deviation from its constant trajectory due to maneuvering.

5-3 Conclusion

This dissertation has examined and expanded the subject of nonlinear-filtering approximation. A new global-filtering approximation procedure has been developed with a particular emphasis on non Gaussian processes. An important practical feature of the proposed method is the method's independence of the model smoothness assumption which is crucial to traditional techniques. Furthermore, a major and equally important byproduct is the generation of a "closed" (in the mean-square-error sense) bilinear model approximation of the original nonlinear system.

The estimation method developed herein has been applied specifically to practical problems to demonstrate its effectiveness and applicability to a variety of a certain class of nonlinear systems. In fact, the degree of ease or difficulty in extending its applicability to general multi-dimensional nonlinear systems is directly related to the ease or difficulty in calculating the multiplicative feedback control-laws, and accordingly in evaluating the required conditional expectation of the nonlinear terms. This is often proven to be difficult, especially for large-scale nonlinear systems. However, for low-order nonlinear systems, this difficulty has been surmounted by a novel decomposition scheme. This scheme alleviates the control problem calculation, and permits a reduction of the multi-dimensional integrations associated with evaluation of the conditional expectation expressions to only a single integration which can then be evaluated analytically. It is noteworthy to point out that a similar approach which is developed in [K3] and has been used in the second example, does not require any control-law calculation but has no global filtering properties either. However, the method requires evaluation of the conditional expectations with corresponding multi-dimensional integrations.

A fundamental limitation imposed on the new approach was the conditionally-Gaussian assumption of x_0 given y_0. This restriction is very basic, because the filtering techniques herein are completely dependent on the choice of the statistical model for the underlying random processes. However, as pointed out earlier, this assumption sometimes may be satisfied under realistic operating conditions, and, of course, it is more general than the traditional Gaussian assumption of both x_t and y_t.

The digital-computer simulation clearly demonstrates the efficiency and filtering accuracy improvement of the TNF over the popular EKF. Thus, the apparent complexity of the algorithm and the slight increase in computation cost might be justified by the significant improvement in the filter performance. But, no claim has been made that this filter is superior to all other existing techniques in all cases. That certainly warrants further investigation.

Finally, the methodological formulation developed in this work is intended to generate further interest and insight into the design of a future filter for general nonlinear systems. Moreover, it is hoped also, that this filter's potential, as an effective filtering algorithm for a nonlinear system, should be fully explored by practical application to communication and tracking.

BIBLIOGRAPHY

A1 Ahmed, N.U., Teo, K.L., "An Existance Theorem on Optimal Control of Partially Observable Diffusions", SIAM J. Control 12, No. 3, Aug 1974, pp. 351-355.

A2 Aidala, V.J., "Behavior of Kalman Filter Applied to Bearings-only Target Motion Tracking," in Advances in Passive Target Tracking, Monterey, CA; Naval Post Graduate School, May 1977, pp. 1-30.

A3 Ames, W.F., Nonlinear Partial Differential Equations in Engineering, Vol. II., Academic Press, New York, 1972.

A4 Astrom, K.J., Introduction to Stochastic Control Theory, Academic Press, New York, 1970.

B1 Bucy, R., Joseph, P., Filtering for Stochastic Process with Applications to Guidance, John Wiley & Sons, New York, 1968.

B2 Benton, S.H., The Hamilton Jacobi Equation: A Global Approach, Academic Press, New York, 1977.

C1 Chou, S.I., "Some Drawbacks of Extended Kalman Filtering in ASW Passive Angle Tracking," in Advances in Passive Target Tracking, Monterey, CA, Naval Post Graduate School, May 1977, pp. 76-113.

D1 Davis, M.H.A., "On the Existance of Optimal Policies in Stochastic Control," SIAM J. Control, II (1973), pp. 587-524.

D2 Davis, M.H.A. Linear Estimation and Stochastic Control, John Wiley & Sons, New York, 1977.

F1 Fleming, W.H., Rishel, R.W. Deterministic and Stochastic Optimal Control, Springer-Verlag, New York, 1975.

F2 Friedman, A. Partial Differential Equation of Parabolic Type, Prentice-Hall, Englewod-Cliff, N.J. 1964.

F3 Friedman, A. "Partial Differential Equations," Holt-Rinehart and Winston New York, 1969.

G1 Gustafson, K. Partial Differential Equations and Hilbert Space Methods, John Wiley & Sons, New York, 1976.

BIBLIOGRAPHY
(Continued)

G2 Gilbert, R. Function Theoretic Methods in Partial Differential Equation, Academic Press, New York, 1969.

J1 Jazwinski, A.H. Stochastic Processes and Filtering Theory, Academic Press, New York, 1970.

K1 Kolodziej, W.J. Conditionally Gaussian Processes in Stochastic Theory, Ph.D. Dissertation, Oregon State University, June (1980).

K2 Kolodziej, W.J., Halawani, T. and Mohler, R.R. "A New Nonlinear Filtering Approximation: Proceedings of the 20th IEEE Conference on Decision and Control, San Diego, CA, Dec. 1981, Vol. II, pp. 798-800.

K3 Kolodziej, W.J. and Mohler, R.R. "Analysis of a New Nonlinear Filter and Tracking Methodology," To appear.

K4 Kushner, H.J. "Approximation to Optimal Nonlinear Filters," IEEE Trans., A.C. Vol AC - 12, pp. 546-556, 1967.

L1 Liptser, R.S. Shiryayev, A.N. Statistics of Random Process I: General Theory, Springer-Verlag, New York, 1977.

L2 Liptser, R.S., Shiryayev, A.N. Statistics of Random Process II, Applications, Springer-Verlag, New York, 1978.

M1 Mohler, R.R. Kolodziej, W.J. "Optimal Control of a Class of Nonlinear Stochastic Systems." IEEE Trans. A.C. Vol AC-26, (1981), No. 5, pp. 1048-1054.

M2 Mohler, R.R. Bilinear Control Process, Academic Press, New York, 1973.

P1 Pryor, C.N. An Adaptive Kalman Filter Tracker For Multi-Mode Range/Doppler Sonar, G. Tacconi (ed) Aspect of Signal Processing. Vol. II, Reidel Publishing Co. Dordrech-Holland, 1977.

P2 Papoulis, A. Probability, Random Variables, and Stochastic Processes, McGraw-Hill Co., New York, 1965.

R1 Rudin, W. Real and Complex Analysis, McGraw-Hill, New York, 1966.

BIBLIOGRAPHY
(Continued)

S1 Stratonovich, R.L. "A New form of Representation of Stochastic Integrals and equations, SIAM J. CONTROL 4 (1966), No. 2, pp. 362-371.

S2 Schwartz, L. and Stear, E.B. "Acomputational Comparison of Several Nonlinear Filters," IEEE Tran on Auto Cont., Feb., (1968), pp. 83-86.

W3 Wong, E. Stochastic Processes in Information and Dynamical Systems, McGraw-Hill, New York, 1971.

APPENDICES

APPENDIX A

WIENER-PROCESS FORMULATION AND WHITE-NOISE FORMULATION EQUIVALENCE

Theorem A1 Levy (L1, Thm 4.1, pp.82) defines a Wiener process as follows:

Let (Ω,F,p) be a probability space and (F_t), $t\varepsilon[0,T]$ be a nondecreasing family of sub-σ-algebras of F. The random process (w_t,F_t), $t\varepsilon[0,T]$, is called a Wiener process if

(i) the trajectories w_t, $t\varepsilon[0,T]$ are Gaussian, continuous (p.a.s.) on $[0,T]$,

(ii) w_t, $t\varepsilon[0,T]$ is a square-integrable martingale [A4] with $w_0=0$, p.a.s. and $E[(w_t-w_s)(W_t-w_s)^*]=(s\wedge t)I$, $t>s$

Thus, any Wiener process is a Brownian motion process [L1],[A4].

Definition A1[A4]

A sequence of quadratic mean square continuous $[x_t^n]$, $t\varepsilon[-\infty,\infty]$ is said to converge to a white noise if for each function f(t), g(t), $\left(\int_{-\infty}^{\infty}|f(t)|dt<\infty\right)$, there exists a positive constant S_0, i.e.

$$\lim_{n\to\infty} E\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(t)\, g(s)\, x_t^n x_s^n dt\, ds\right) = S_0 \int_{-\infty}^{\infty} f(t)\, g(t)\, dt.$$

As pointed out earlier, a stochastic differential equation with an additive, exciting, white-Gaussian noise is given by

$$\frac{dx}{dt} = f(x_t,t) + G(x_t,t)\,\beta_t \quad . \tag{A-1}$$

Here, β_t is a white noise, and is thus neither mean square Riemann integrable nor integrable with probability one. Hence equation (A-1) is not mathematically meaningful as it stands. However, if the white noise is as defined in (A1) and is considered as the formal derivative of the Wiener process (w_t, $t\varepsilon T$), (Thm.A1), then (A-1) may be considered formally equivalent to

$$dx_t = f(x_t,t)dt + G(x_t,t)dw_t, \quad t\varepsilon T. \tag{A-2}$$

At least formally it is known that $\int_0^t \beta_t ds$ has all the properties of Brownian motion, w_t. Hence, (A-1) can be made meaningful in terms of a stochastic integral equation [D2].

$$x_t = x_0 + \int_0^T f(x_s,s)ds + \int_0^T G(x_s,s)dw_s \tag{A-3}$$

The last integral in (A-3) is interpreted as a stochastic integral which needs to be defined. Since, as it may be recalled, w_t has a realization of unbounded variation in any small interval of time, the last-integral cannot be defined in the usual Lebesque-Stieltjes sense. One generally accepted definition is due to Ito [L1] and is often referred to as an Ito stochastic integral, and (A-2) is called the Ito stochastic differential equation interpreted in terms of the Ito calculus [A4] which is not compatible with results of ordinary calculus. Accordingly, the following definition is appropriate.

Definition A2 [L1], [K1].

A process x_t is said to satisfy (A-2) for $t\varepsilon[0,T]$, with initial state x_0 if

(i) for all $t\varepsilon[0,T]$, $\int_0^t G(x_s,s)dw_s$ can be interpreted as a stochastic integral,

(ii) for all $t\varepsilon[0,T]$, x_t is almost surely equal to the random variable

$$x_0 + \int_0^t f(x_s,s)ds + \int_0^t G(x_s,s)dw_s.$$

Under certain conditions imposed on x_0,f,G [L1], (A-2) has a unique, strong/weak, sample-continuous, Markov solution. The strong-solution notion is defined as follows

Definition A3 [L1]

For a given complete probability space, (Ω,ζ,p), and a Wiener process w_t, the stochastic differential equation (A-2) has a strong solution x_t if:

$$1)\quad P(\int_0^t ||f(x_t,t)||dt<\infty) = 1 \qquad \text{(A-4)}$$

2) $P(\int_0^T ||G(x_t,t)||^2 \, dt<\infty) = 1$ (A-5)

3) Condition (ii) of Definition A2 is satisfied. (A-6)

That implies that $\zeta_t^X \leqslant \zeta_t^W$,

where $\zeta_t^X = \sigma$-algebra $[x_s;\ 0 \leqslant s \leqslant t]$,

$\zeta_t^W = \sigma$-algrebra $[w_s;\ 0 \leqslant s \leqslant t]$.

The weak solution notion is defined as:

Definition A4 [L1].

Let $F(\alpha) = P(n \leqslant \alpha)$, probability measure for n random variables. If there exist (Ω,ζ,p), x_t, p-a.s. continuous, conditions (A-4), (A-5), (A-6) are satisfied, and $F(\alpha = P(x_0 \leqslant \alpha_0)$. Then x_t is a weak solution to (A-2) which implies $\zeta_t^X \geqslant \zeta_t^W$, and ζ_t^X, ζ_t^W as defined previously.

It is noteworthy to point out that the use of Ito calculus, which led to the definition of the Ito integral $\int_0^t G(x_s,s)dw_s$ in equation (A-3), results in adding a correction term to the result of the ordinary differentiation rules when stochastic process is differentiated. This will help in transforming the differential dx_t as in equation (A-2), into a form that can hopefully be recognized as the differential of a known function. For example, consider the following stochastic differential equation.

$$dx = xdw, \tag{A-7}$$

with initial conditions $x(0) = 1$, $w(0) = 0$, and $E(dw(t)^2) = dt$. Using ordinary integration rules or Stratonovich rule, the analytical solution is

$$x(t) = \text{Exp } w(t) \ . \tag{A-8}$$

However, using Ito integral, the solution is

$$x(t) = \text{Exp } [w(t) - 0.5t], \tag{A-9}$$

which is actually the integral of the following stochastic differential equation

$$dx = x \, dw + 0.5 \, x \, (dw)^2, \tag{A-10}$$

where $(dw)^2 = dt$.

The last term in equation (A-10) is what is referred to as the correction term. Thus, the addition of the correction term to equation (A-7) will make it compatible with ordinary rules of differentiation and integration.

In simulating equation (A-3) or its differential equivalence (equation (A-2)) special care must be given to the simulation of the term $\int B(x_s,s)dw_s$ because (dw) is highly uncorrelated. Thus, it is difficult to find an appropriate integration step-size Δt. However, the following approximation scheme is appropriate.

Let R be the maximum rate of change and defined as

$$R = \max_{x,t} \left[\frac{|F(x,t)|}{1 + |x|} + \frac{G^2(x,t)}{1 + x^2}\right] . \tag{A-11}$$

Now let $\Delta t \ll \frac{1}{R}$, $E\,(dw)^2 \simeq \Delta t$.

Then, equation (A-2) or equation (A-3) can be integrated using digital computer as follows

$$\begin{aligned} x_{k+1} = x_k &+ \Delta t\, F(x_k, t_k) + \Delta w\, G(x_k, t_k) \\ &+ 0.5\, \Delta t\, G(x_k, t_k) \frac{\partial G(x_k,t_k)}{\partial x} . \end{aligned} \tag{A-12}$$

Hence, the last term in equation (A-12) is the correction term which is essential for the simulation of the Ito differential or (integral) equation by digital-computer.

Let us return for a moment to equation (A-3) and try to find some physical interpretation to it. If $x(t)$ is the state of a dynamical system, then the terms on the right-hand side of equation (A-3) can have a nice interpretation. The term x_0 is just the initial condition. The first integral describes the evolution of the component of the state with time. The second integral can be considered as the irregular component, which is entirely due to noise. Furthermore, it is well known, (using practical engineering assumptions) that any continuous noise can be regarded as a smooth transformation of a standard Wiener process. Unfortunately, the standard Wiener process is not differentiable with respect to time, hence no such dw exists. Thus, only the stochastic integral as in equation (A-3) is available for modeling real systems. However, as pointed out earlier, if the white Gaussian noise is considered as a formal derivative of the Wiener process then the stochastic differential equations (A-1), (A-2) would have a particular appeal to engineering applications. Furthermore, the white-noise concept allows us to manipulate the Wiener integral (which is a special case of the Ito integral, if certain conditions about the function $G(x_s,s)$ are satisfied i.e. $G(x_s,s)$ is square integrable) as an ordinary integral but not in the same sense as the Stieltjes integral is defined. It also allows a suitable mathematically tractable model for many continuous physical noises encountered in real engineering systems.

Now that equations (A-1), (A-2) and (A-3) have been presented formally, let us discuss briefly the use of such equations in modeling real physical processes if they are going to be of any practical value. As it is known in many practical problems both in control and communications, differential equations arise from the laws of nature, but it is not advisable to take derivatives of certain signals. Thus, it is appropriate to model these signals by so-called stochastic differential equations. In addition, almost any mathematical model of a physical process involves a degree of idealization that produces a good match with reality only within certain ranges of the parameters involved. Thus, the result of such a modeling can be judged only by comparison with practical experiments within the prescribed ranges of the parameters involved. For example, the erratic motion of a particle or point mass submerged in a fluid caused by impact of the molecules of the liquid on the particle. The force acting on the particle can be approximated by

$$-\alpha x(t) + \beta\, u_n(t). \tag{A-13}$$

where x(t) denotes a position of the velocity of the particle at time t. Here, the first term in (A-13) represents friction or drag, while the second represents the push imparted upon the particle by some projecting force which is random in nature. If the process $u_n(t)$ is replaced by white Gaussian noise n_t. Then the motion of the particle can be approximated by

$$\dot{x}(t) + \frac{\alpha}{m}\, x(t) = \frac{\beta}{m}\, n_t, \quad t \geqslant 0. \tag{A-14}$$

Here m is the mass of the particle, and $\dot{x}(t)$ is the acceleration. Hence we arrived at the famous Langevin equation. Thus, the velocity coordinate, (with certain assumptions about the initial conditions), will be an Ornstein-Uhlenbeck process which can be obtained as the solution of the stochastic equation

$$dx_t = -\alpha x_t + c dw_t, \qquad \text{(A-15)}$$

where w_t is a Wiener process and α,c are positive constants.

The relevance of the model (A-14) or equivalently (A-15) can now be tested, for example, by observing to motion of the particle and deciding whether the statistic, (mean, variance), of the displacements of the particle can be described as a white Gaussian noise.

The practical implementation of the concept of weak/strong solutions depends in large on the particular application at hand, and the system modeling approach. A strong solution usually deals with a "given" Wiener process as a model for a wide spectrum random noise, while a weak solution is based on the promise that there exists a Wiener process which can be used as such a model. Thus, if the physical properties of a given problem specifies the probability space (Ω,ζ,p), the system's possible outcome event set(ζ_t), $t\varepsilon[0,T]$, and the Wiener process $w = (w_t)$ then the strong solution approach is appropriate. On the other hand, if the physical nature of the problem does not specify the complete probability space (Ω,ζ,p), then, the weak solution approach might be more suitable as a modeling approach. Thus, we may construct a probability space (Ω,ζ,p), a system (ζ_t), $t\varepsilon[0,T]$, and a Wiener process $w=(w_t)$, for which (A-5)-(A-6) are satisfied (p.a.s.), to satisfy the modeling purpose.

APPENDIX B

AUXILIARY RESULTS

Lemma B1 [D2]

Let $[x_t^n]$ and $[y_t^n]$ be two jointly normal processes and let $E(x/y_t) = \bar{x} = P_t Y_t$, the projection of x onto H_t^y. Then, the conditional characteristic function, (for the vector case),

$$\Psi_{x/y}(\bar{w}) = \mathrm{Exp}[j\bar{w}\bar{x}_t - 0.5 \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j u_{ij}], \tag{B-1}$$

where $\bar{w} = (w_1, w_2, \ldots w_n)$,

$\bar{x}_t = (\bar{x}_1, \bar{x}_2, \ldots \bar{x}_m)$, the conditional expectation,

$u_{ij} = u_{ji} = E((x_i - \bar{x}_i)(x_j - \bar{x}_j)^* / y_s;\ s \leqslant t)$,

the conditonal covariance.

The proof is an extension of the proof given in [D2, pp. 53] for the scalar case, and is omitted for brevity.

Lemma B2, [P2, pp. 146]

Let $[x_i]$, i=1,2,...n be jointly normal processes with the conditional characteristic function $\Psi_{x/y}(\bar{w})$ as given in Lemma B2.

Then, the conditonal expectation of order $r = k_1 + k_2 + \ldots k_n$, is

$$J^r E[(x_1, \ldots x_n / y_s;\ s \leqslant t)] = \left.\frac{\partial^r \Psi_{x/y}(w_1, w_2 \ldots w_n)}{\partial w_1^{k_1} \partial w_2^{k_2} \ldots \partial w_n^{kn}}\right|_{\bar{w}_n = 0} \qquad \text{(B-2)}$$

The proof is parallel to the proof given in [P2] using the above two lemmas; the following results are noted;

$$E(x_1 x_2 x_3 / y_s;\ s \leqslant t) = \bar{x}_1 u_{23} + \bar{x}_2 u_{13} + \bar{x}_3 u_{12} + \bar{x}_1 \bar{x}_2 \bar{x}_3 \quad , \qquad \text{(B-13)}$$

$$E(x_1 x_2 x_3 x_4 / y_s;\ s \leqslant t) = u_{12} u_{34} + u_{13} u_{24} + u_{14} u_{23} + \bar{x}_1 \bar{x}_2 u_{34}$$

$$+ \bar{x}_1 \bar{x}_3 u_{24} + \bar{x}_1 \bar{x}_4 u_{23} + \bar{x}_2 \bar{x}_3 u_{14} +$$

$$\bar{x}_2 \bar{x}_4 u_{13} + \bar{x}_3 \bar{x}_4 u_{14} + \bar{x}_1 \bar{x}_2 \bar{x}_3 \bar{x}_4 . \qquad \text{(B-14)}$$

$$E(x_1x_2x_3x_4x_5/y_s;\ s\leqslant t) = \bar{x}_1\bar{x}_2\bar{x}_3\bar{x}_4\bar{x}_5 + \bar{x}_1[u_{23}u_{45} + u_{24}u_{35} + u_{25}u_{34}]$$

$$+ \bar{x}_2[u_{13}u_{45} + u_{14}u_{35} + u_{15}u_{34}] + \bar{x}_3[u_{15}u_{24} + u_{25}u_{14}$$

$$+ u_{12}u_{45}] + \bar{x}_4[u_{12}u_{35} + u_{13}u_{25} + u_{15}u_{23}]$$

$$+ \bar{x}_5[u_{12}u_{34} + u_{13}u_{24} + u_{14}u_{23}]$$

$$+ \bar{x}_1\bar{x}_2\bar{x}_3u_{45} + \bar{x}_1\bar{x}_2\bar{x}_4u_{35} + \bar{x}_1\bar{x}_2\bar{x}_5u_{34}$$

$$+ \bar{x}_2\bar{x}_3\bar{x}_4u_{15} + \bar{x}_2\bar{x}_3\bar{x}_5u_{14} + \bar{x}_3\bar{x}_4\bar{x}_5u_{12}$$

$$+ \bar{x}_1\bar{x}_4\bar{x}_5u_{23} + \bar{x}_1\bar{x}_3\bar{x}_5u_{24} + \bar{x}_2\bar{x}_4\bar{x}_5u_{13}$$

$$+ \bar{x}_1\bar{x}_3\bar{x}_4u_{25} \qquad \text{(B-5)}$$

Here $u_{ij} = E\{(x_i-\bar{x}_i)(x_j-\bar{x}_j)^*/y_s;\ s\leqslant t)\}$ = the condition covariance.