MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

UNITED STATES AIR FORCE

AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOG

Wright-Patterson Air Force Base,Ohio

79 01 30 13

FUNDAMENTAL LIMITATIONS

OF OPTICAL TRACKERS

THESIS

AFIT/GEO/EE/78-4   John M. Santiago, Jr.
                     2nd Lt   USAF

Approved for public release; distribution unlimited

79 01 30 135

AFIT/GEO/EE/78-4

FUNDAMENTAL LIMITATIONS
OF OPTICAL TRACKERS.

THESIS Master's thesis,

Presented to the Faculty of the School of Engineering

of the Air Force Institute of Technology

Air Training Command

177 P' in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

by

John M. Santiago, Jr

2nd Lt USAF

Graduate Electro-optics

December 1978

Approved for public release; distribution unlimited

Preface

This thesis was under the sponsorship of the Air Force
Weapons Laboratory at Kirtland Air Force Base in New Mexico.
It is a theoretical study to determine the performance capa-
bilities of optical trackers when fundamentally limited by the
quantum nature of light and the photon detectors used to detect
the optical fields. This study involves a knowledge of optical
communications, optimal filter theory, and software simulation
techniques that made the study interesting.

I would like to first thank Capt Stan Z. Lewantowitz for
sponsoring this thesis, and Maj Joseph W. Carl for his interest
in this work. I am extremely thankful to my thesis advisors,
Dr. Peter S. Maybeck and Capt Stanley R. Robinson, for their
guidance, expertise, and patience that they have given to
me during this research study. Also, I would like to give many
thanks to Cindy Held and Sheri Vogel for burning the midnight
oil in typing this final product. Finally, even though thousands
of miles away, I would like to thank my mom and dad and the rest
of my family for their encouragement and moral support during this
period of learning.

<div align="right">John M. Santiago, Jr.</div>

Contents

List of Figures

List of Tables

Abstract

A problem is considered to determine the tracking capabilities of an estimator, applied in optical sensing. The estimator tracks the centroid of a one-dimensional Gaussian-shaped intensity based on time-space point process measurements. The centroid assumes to move dynamically as a First Order Gauss-Markov process. Filter performance is described by steady-state upper and lower bounds on mean-square-error (MSE) which are evaluated as a function of two physically motivated parameters: average number of photons detected in a coherence time of centroid dynamics and mean square value of centroid jitter normalized by the square of the beam width. The parameters establish regions of operations where upper and lower bounds converge to the actual MSE. Results from the Monte Carlo simulation demonstrates the bounds' usefulness. Noise measurements from dark current or background radiation are included in the simulation. Results show that the filter is very sensitive to these measurements, resulting in very poor tracking. Ad hoc methods of filter tuning and residual monitoring are employed to improve tracking performance; results indicate that filter performance can be improved substantially through residual monitoring.

Fundamental Limitations
of Optical Trackers

I. Introduction

Alignment problems exist in any optical system. To image
a point object over large distances requires aligning an optical
receiver with the object. When the object is within the receiver's
field-of-view, an image is formed on the photodetector surface.
This is shown in Fig. (1). Because the optical receiver has a
relatively narrow field-of-view, aligning the receiver becomes
difficult. Furthermore, because of variations of index of
refraction in the atmosphere and other outside disturbances
that cause. relative motion between receiver and object, active
alignment is necessary. Thus, optical trackers are needed to
keep the object within the receiver's field-of-view.

Background

Several methods are currently being developed in the Air
Force to track the target in presence of several disturbances.
The disturbances include target motion, mirror vibration, beam
jitter, stochastic effects of the atmosphere, and other processes
that can cause relative motion between the beam and target.

Method of Tracking.

One scheme to accomplish this tracking
task uses a direct detection optical receiver to determine the
arrival angle of the incident radiation at the receiver. The
arrival angle is converted (through optics) into a position of
a spot of light on the active surface of a photodetector array.
Once this angle arrival is determined, it can be used to
generate error signals to other tracking components to correct
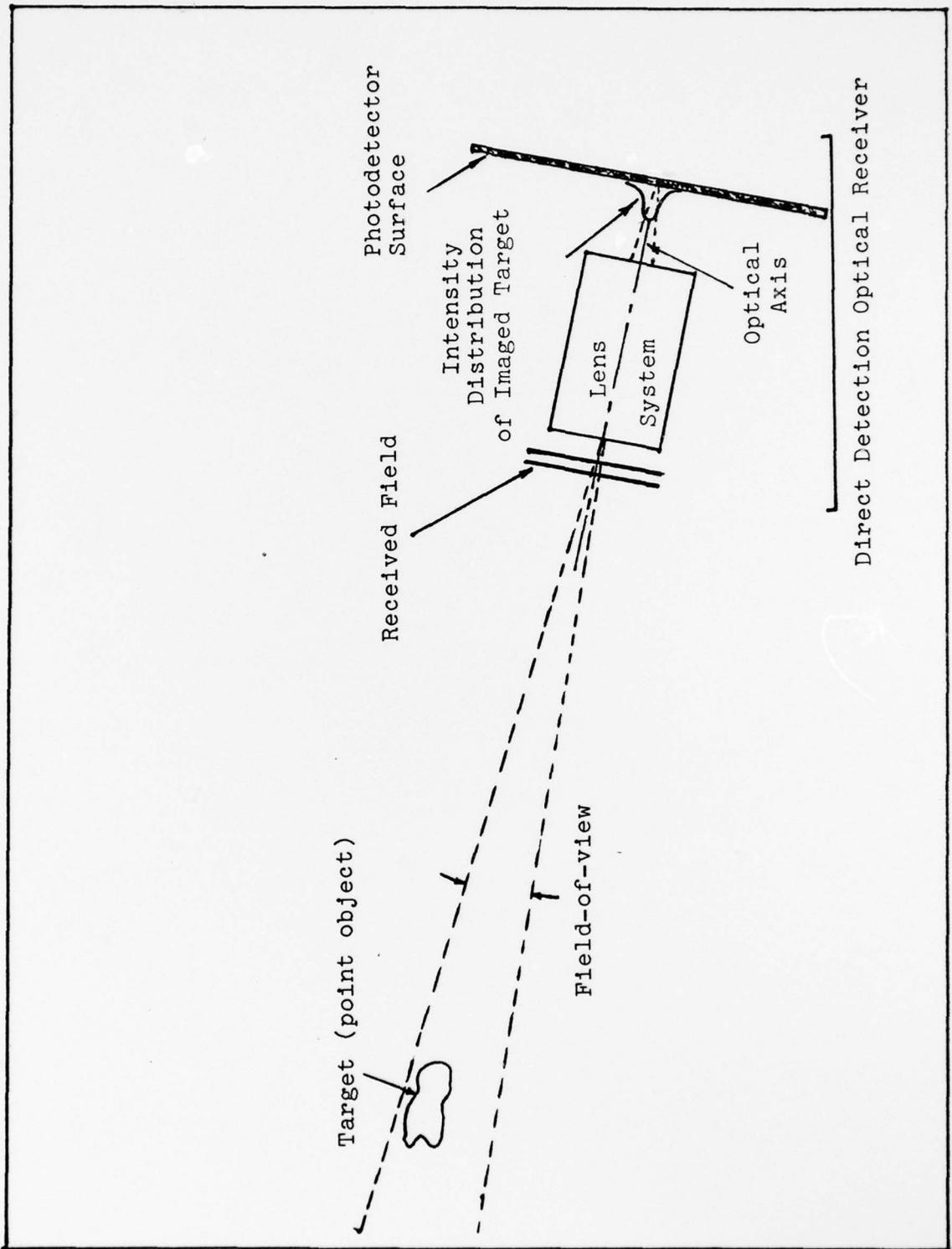
Figure 1. Target Within Field-of-View of Direct Detection
Optical Receiver

2

any offset alignment between the receiver and target that may occur. Because of disturbances that cause relative motion between the beam and target, the beam's arrival angle varies at the optical receiver. Thus, variations of arrival angle are therefore converted into variations of the spot's position on the detector surface. The spot of light moves dynamically on the surface, releasing photoelectrons statistically distributed about the spot's position of maximum intensity.

A minimum-mean-square estimator which will be known as the Snyder Filter estimates the position of maximum intensity based on the observations of the photoelectrons (Ref. 1). These photoelectrons are measurement data to the filter in order to obtain an estimate of the spot's current position. Here, the Snyder Filter assumes that the measurements only include the target's signal intensity. However, other disturbances also cause photoelectrons to be released from the surface. These disturbances which degrade the Snyder Filter's tracking capability are caused by background radiation from the sky, moon, sun, stars, and other sources, and dark current present in the absence of photoexcitation on the photodetector. Each of the noise sources are discussed in the next chapter. It should be noted, however, that thermal noise is ignored for photodetectors with high gain, such as a photomultiplier or avalanche photodiode.

Quantum Effects. At higher frequences, the quantum nature of light is an important consideration in many optical measurement systems which use photon or quantum detectors to detect these optical fields. In this case, light can be described as

a collection of discrete packets of energy called photons. With a given probability, each photon converts into an electron. These photoelectrons, distributed in time and space, are released from the photodetector surface. Because of these quantum effects in the detection process, especially when the field intensity is low at the receiver's input, there is an uncertainty associated with estimating the spot's position of maximum intensity based on the sensed photoelectrons. Theoretically, both the quantum nature of light and the nature of the detection process fundamentally limit the tracking capability of optical trackers (in other words, the ability to detect the signal photons is the fundamental limit).

Because of the quantum effects associated in the detection process, the statistics of the receiver's output process must be described accurately; the Snyder Filter is based on these statistics. Experimental evidence indicates that the statistics of the output can be modelled as a random point process in time and space conditioned on knowing the optical field incident on the receiver (Ref. 2). A random point process by definition is "a mathematical model to describe a physical phenomenon characterized by highly localized events distributed randomly in a continuum (Ref. 3:2)." The discrete events are the photo-electrons or photoconversions distributed randomly on the photo-detector surface. Associated with each event are the temporal and spatial location of the photoelectron. Here, the event location is used as distributed measurement or data,

in order to estimate the current position of maximum intensity. To summarize, the point process model describes the fundamental uncertainty of photon-to-electron conversion in the measurement or detection process.

Measurement Models (Gaussian versus Point Process Descriptions). To contrast the point process measurement description with other noise models, several Gaussian descriptions modelling different types of noise are briefly presented. A stationary white noise is one model which describes noise as having a power spectral density of infinite bandwidth, having energy at all frequencies. Also, infinite bandwidth implies that the noise samples are uncorrelated in time. Although no noise has infinite bandwidth, the white noise model does provide a reasonable description for wideband noise having a wider bandpass than the physical system of interest. In other words, the wideband noise appears white with respect to the system. In many instances, however, noises are not well modelled as uncorrelated in time (e.g., bandlimited noise), and other models must be exploited. For such causes, a Gaussian white noise, driving the input of a small linear system, can generate an output that duplicates or closely approximates the second order characteristics of a given noise source. The linear system, or "shaping filter", duplicates the second order statistics of the actual noise by generating a Gaussian process with the same second order statistics. For example, the statistics of a zero-mean exponentially time-correlated (First Order Markov) noise can be duplicated by taking the output

5

statistics of a first order lag driven by white-Gaussian noise of zero mean. This model approximates a variety of experimentally observed random behavior that have wide or narrow bandwidths with a flat power spectral density. Another noise model is the Second Order Markov process model which describes phenomena having vibration, bending or any other periodic random characteristics. This process can be generated as the output of a second order system driven by white noise.

The different models that were just described typically model noise as due to a number of sources that are always present (or continuous) in a dynamics system or measurement device. By the Central Limit Theorem of Statistics, when a number of random variables are added together, the sum of these random variables has a probability distribution which is nearly Gaussian, regardless of the shape of the individual densities. The Gaussian models are exploited in many cases where macroscoptic phenomenon are observed. On the other hand, the quantum effects are observed as a microscopic interaction between the photons in the incident field and the atoms on the photodetector surface. The result of this discrete interaction is the probability of releasing a photoelectron at a particular time and location from the photodetector array. Thus, various disturbances affecting a system or measurement device cannot always be described as having additive corrupted noise with Gaussian statistics. To incorporate the quantum effects or uncertainty, the random point process description is used to model the measurement process from the optical receiver under limited operations.

Assumptions of the Snyder Filter. This report assumes
that the point object is within the field-of-view of the
optical receiver and is imaged as a single spot on the photo-
detector array. In other words, the tracker has acquired the
target and, hence, the tracker is said to be in "fine-tracking"
mode. As mentioned previously, the solution of estimating the
position of maximum intensity based on released photoelectrons
is the Snyder Filter. For this research, the Snyder Filter is
implemented in open-loop configuration (i.e., no control or
feedback input). The Snyder Filter models the spot as having
a Gaussian-shaped intensity profile and the spot drifts
dynamically as a First Order Gauss-Markov process. These
characteristics and the Snyder Filter are discussed in detail
in the next chapter.

In addition to the above assumptions, this estimator
assumes its measurements are from a photodetector array capable
of measuring the "exact" temporal and spatial component of the
photoelectron. This idealization places a limit in tracking
capabilities which can be accomplished with photodetectors
arrays of finite resolution, such as a quadrant photomultiplier.
However, the Snyder Filter assumes that its measurement data
contain no extraneous (or noise) measurements. That is, photo-
electrons are not released due to background radiation or dark
current. This report investigates the tracking performance of
the Snyder Filter based on observations of the sensed photo-
electrons as input data.

7

Problem

The problem is to establish the performance capabilities of the Snyder Filter. In this case, the issue is whether this minimum mean-square-error estimator commits errors small enough to yield good tracking. Thus, the problem is to describe the error performance of the Snyder Filter in order to make this evaluation. Once the performance description is established (under the assumptions of no background or dark current), an additional objective is to determine the sensitivity of the Snyder Filter when noise data is presented to the filter as input. To improve filter performance in the presence of the above disturbances, ad hoc methods which alter the filter structure will be investigated.

Overview and Approach

Chapter II describes the models employed in the tracking scheme and establishes conditions in tracking performance under which the Snyder Filter is evaluated. The chapter begins by describing the model of the direct detection optical receiver and its detection process. In this section, the semiclassical approach is used to describe the photodetection process. The results of this analysis are used to derive the statistics of the output process from the optical receiver (Appendix A). Following this discussion is a description of the detector noise which affects the output of the receiver.

Following the model of the optical receiver is a section on the Snyder Filter. A characterization of the spot's

8

intensity profile and its dynamics are given. Following the
subsection is a portrayal of position-sensitive measurements
available to the Snyder Filter. As noted earlier, these
measurements consist of point process photoconversions in time
and space. Concluding the section are the filter equations
of the Snyder Filter.

Chapter II finally concludes with a description of
performance parameters to describe the mean-square-error
(MSE) of the Snyder Filter. In this case, performance
descriptions of estimators and controllers that employ point
process observations in time and space are available in recent
literature (Ref. 4). However, this report addresses only the
estimation performance of the Snyder Filter. Also, the
Snyder Filter is a minimum mean-square-error (MMSE) estimator,
under ideal conditions (i.e., input data from background
radiation or dark current are not included). The MSE, also
denoted as $E[P(t)]$, is used to determine whether the errors
commited by the Snyder Filter are sufficiently small for
tracking purposes. However, the MSE cannot be evaluated in
closed form; thus, upper and lower bounds are derived rather
than solving for MSE. In Appendix B, the models and conditions
of the Snyder Filter are applied in the upper and lower bounds.

In the appendix, two physically motivated parameters are
defined from these bounds to establish conditions under which
the Snyder Filter yeilds good tracking: the average number
of photoelectrons in a coherence time of the spot's
dynamics and the mean square value of the spot's jitter

9

normalized to the square of the beamwidth. These parameters
establish regions of operations where the upper and lower
bounds converge to the actual MSE (i.e., bounds are identical).
The parameters are discussed in detail at the end of
Chapter II. Note that the performance measures themselves
assume perfect data; that is, these measures of performance
are not valid for cases which include data from dark current
and background radiation.

To demonstrate the tracking performance provided by the
above parameters and to determine the filter's sensitivity
to noise measurements, a Monte Carlo simulation and performance
analysis is performed on a digital computer. Chapter III develops
algorithms to simulate the models in the computer required to
perform this study. The chapter begins with a section on simu-
lating the spot's dynamics on the detector surface and the
required noise generators are discussed next. Following these
sections are time-space point process generators. Two methods
are used to generate these data measurements: one, the Binomial
Method, uses the incremental definition of the point-process,
and the other, the Conditional Method, involves Bayes' rule
and the statistics of the output of the optical receiver. The
Binomial Method involves a threshold test to declare an event
(photoconversion) at each infinitesimal area on the detector
surface. The Conditional Method calls a uniform random variable
distributed between zero and one from the computer system and
transforms it into another random variable with a desired
probability distribution. This transformed variable is a

realization of either the time or space coordinates associated with the photoconversion process. After describing these two methods, the theoretical and practical aspects of implementing each method in the computer are discussed. Concluding Chapter III is the performance analysis which characterizes the error process generated by the Snyder Filter. Given in this section are the sample statistics, plots, and test cases required to evaluate filter performance. Assuming perfect measurements (i.e., no background or dark current events), the actual performance of the Snyder Filter is compared with the upper and lower bounds on MSE. This analysis should indicate where the actual performance lies with respect to the bounds. Then noise measurements are included in the analysis to determine the filter sensitivity to these measurements. Finally, ad hoc methods of residual monitoring and filter tuning are employed to enhance tracking performance in the presence of the noise measurements.

Chapter IV contains results from the simulation. For perfect measurements, the simulation results indicate the usefulness of the bounds in describing the actual performance of the Snyder Filter. Also, these results show for low signal events per coherence time that the upper bound describes the actual performance. When measurements from background radiation or dark current are included in the simulation, the estimator was found to be quite sensitive to these noise measurements, resulting in very poor tracking. Finally, it was found that

tracking performance can be improved substantially by an
ad hoc procedure called residual monitoring.

The last chapter is a summary of conclusions and
recommendations for further research.

II. The Snyder Filter and Its Open Loop Estimation Performance

The Snyder Filter is a recursive data processing algorithm. This algorithm employs point process observations in time and space to estimate optimally the position of maximum intensity on a detector array of infinitesimally fine partition. However, its estimate is only as good as the model describing the output statistics of the photodetector array (the key component in the optical receiver). Therefore, the model of the direct detection optical receiver and its detection process deserves description. The modelling will involve a statistical description of the receiver's output. Following the receiver model, the Snyder Filter is discussed. Concluding the chapter is the tracking performance by which this estimator is to be judged using upper and lower bounds on MSE. For this case, two parameters are defined to describe the actual performance of the Snyder Filter. The concepts are presented in one spatial dimension.

Optical Receiver Model

There are two classes of optical receivers, direct detection receivers and heterodyning receivers. The difference between these two receivers is that the former operates to detect the instantaneous power in the collected field at receiver's input and the latter operates to detect a combined field formed by mixing a locally generated field with the received field. This study is concerned with direct

detection receivers.

General Description. The direct detection receiver is
also known as a noncoherent detector which is shown in Fig.
2. The receiver consists of a lens system and a photo-
detector. The lens system, also referred to as the front
end of the receiver, performs two functions. One is to
focus the optical field onto the photodetector surface at
a focal point. Another function is to provide some degree of
reducing the intensity of the background noise (i.e., radia-
tion from the sun or sky) prior to photodetection. The photo-
detector then converts the incident field into an electrical
signal.

Photodetectors. Photodetectors consist of two types:
photon and thermal detectors. Because thermal detectors have
a slower response for the frequencies of interest, they are
not considered further (Ref. 5:87). On the other hand, photon
or quantom detectors respond quickly at the optical frequencies.
An excellent discussion on quantum detectors, their description,
performance, and characteristics, is given in Refs. 5:87 and
6:298 . However, of all the available types of detectors using
photosensitive materials, all behave according to quantum-
mechanical principles (Ref. 2:6). Figure 3 portrays the photo-
detector model. These detectors use photosensitive materials
to produce electrical signals in response to changes in the
input field intensity. The signals are electrons, released
from the photodetector surface, and they are collected at
the anode by an electric field. This collection and movement

14

Lens
Spatial
Filter

Photodetector
Array

Received
Field

Focused
Field

Frequency
Filter
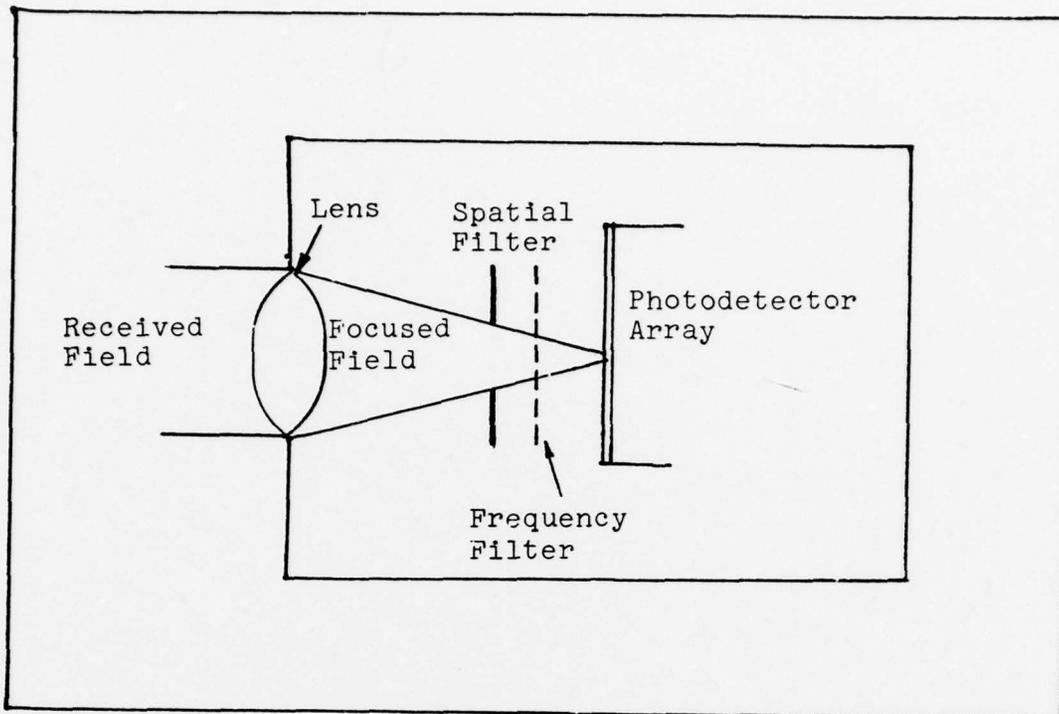
Figure 2.   Direct Detection Optical Receiver
(copied from Ref. 2:5)

of charge at the anode creates an output current. Because

the conversion of optical field to electron flow and the

arrival times of electrons at the anode are probabilistic, the

observed output current is a random process in time and space

(Ref. 2:39). This fact is true even if the input field is

deterministic. In short, the detection operation inherently

induces a randomness in the detector output.

To model the relationship between the optical field and

the number of photoelectrons released involves a treatment

in quantum mechanics (Ref. 3). This treatment has two

approaches to describe the above phenomenon. One is to

describe the field as a collection of discrete packets of

energy called photons. Here, each photon produces an electron
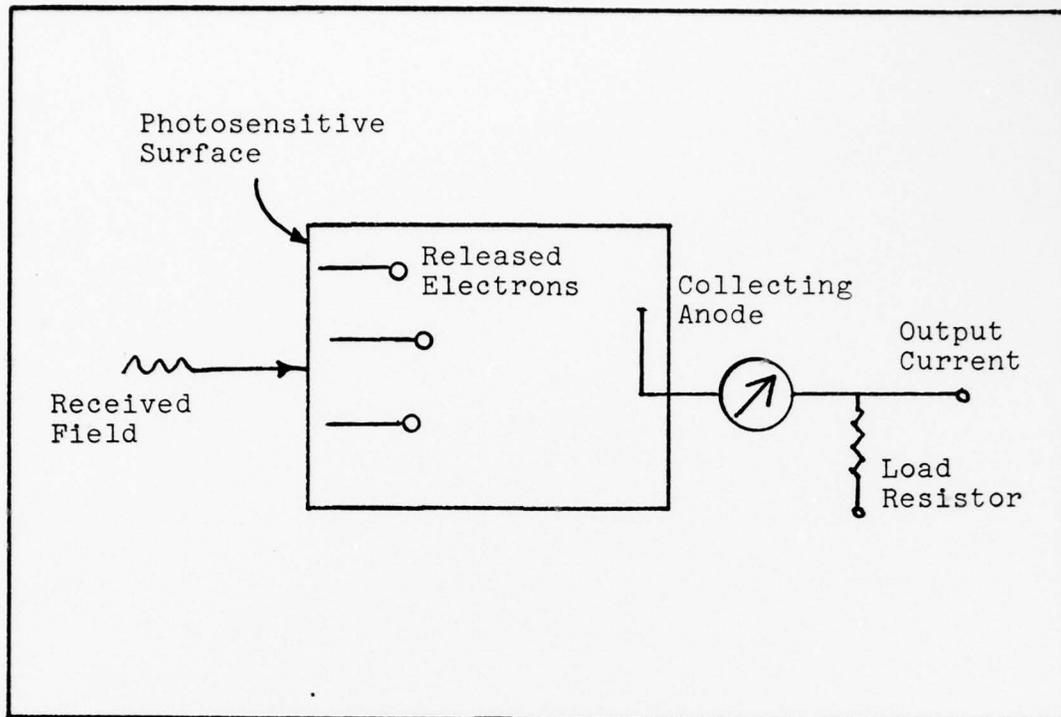
Figure 3. Photodetector Model
(copied from Ref. 2:38)

at the detector surface with a given probability. This

probability of conversion is termed the quantum efficiency,

η. The second method, referred to as a semiclassical analysis

(a direct result of quantum-mechanical considerations), treats

the focussed field as a wave, and it incorporates a statistical

relation to describe the interaction between the optical wave

and the detector surface. The second method is preferred for

its simplicity and the added insight it provides. An outline

of the second approach is found in Ref. 2 and a detailed

description is given in Ref. 7 . Since a detailed description

of the semiclassical analysis is quite lengthy, only results

of this semiclassical approach are given. These results are

16

as follows: in a short infinitesimal time interval, Δt, the release of an electron from an atom at a differential lenght $\Delta \ell$ on the detector is proprotional to the field intensity over the observation area, $\Delta \ell \Delta t$. That is,

$$P \begin{bmatrix} \text{one electron released} \\ \text{in a length } \Delta \ell \text{ during} \\ \text{the time } \Delta t \end{bmatrix}_{\Delta \ell \Delta t \to 0} \simeq \gamma I(t,r) \Delta t \Delta \ell \quad (1)$$

where $P[\cdot]$ denotes probability of, γ is a proportionality constant, $I(t,r)$ denotes the field intensity at time t and point r on the detector surface, and $\Delta \ell$ is located at r. Other results from this approach are as follows:

$$P \begin{bmatrix} \text{no electrons released} \\ \text{from } \Delta \ell \text{ during } \Delta t \end{bmatrix}_{\Delta \ell \Delta t \to 0} \simeq 1 - \gamma I(t,r) \Delta t \Delta \ell \quad (2)$$

$$P \begin{bmatrix} \text{more than one electron} \\ \text{released from } \Delta \ell \\ \text{during } \Delta t \end{bmatrix}_{\Delta \ell \Delta t \to 0} \simeq 0 \quad (3)$$

Along with these results is an assumption that the release of electrons from disjoint observation areas $\Delta t \Delta \ell$ are treated as independent events. This assumption and Eqs (1) through (3) mathematically model the photodetecting surface.

Detector Noise. With no noise sources, the output current is given by

$$i(t,r) = q \sum_{k=1}^{n} \delta(t-t_k) \delta(r-r_k) \quad (4)$$

where q is the electronic charge, t_k is the kth event
arrival time, r_k is the corresponding event location, r
denotes a position on the detector surface, n is the number
of events in an observation area, and $\delta(t)$ and $\delta(r)$ are
delta functions. Eq. (4) assumes that the detector has infinite
bandwidth (e.g., a photomultiplier or avalanche photodiode
fits closely to this description). The relation also indicates
that the current is a random process since n, t_k, and r_k are
random variables.

However, the imperfection of the optical receiver and
background radiation also affect the output current. In
general, the output current from the photodetector array
at (t,r) can be represented as

$$i(t,r) = i_s(t,r) + i_b(t,r) + i_d(t,r) + i_{th}(t,r) \quad (5)$$

where $i_s(t,r)$ is due to signal current, $i_b(t,r)$ is current
due to background noise, $i_d(t,r)$ is due to dark current, and
$i_{th}(t,r)$ is current due to thermal noise. The dark current
is a result of current flowing in absence of photoexcitation.
Also, the dark current is well modelled as a point process
(Ref. 5:148) with a rate function that is constant (i.e.,
$\lambda_d(t,r) = \lambda_d$ is termed homogeneous(Ref. 3:54)). Further,
the background noise can be associated with a point process
in nighttime operation, and its rate function is homogeneous
provided that three conditions are satisfied: (1) background
electric field has zero mean, (2) background intensity is
constant, and (3) the product between the ensemble average

18

of the number of electrons, n_b, and the bandwidth of the background radiation, $\frac{1}{\tau_b}$, is much less than 1. The last condition, $n_b \tau_b \ll 1$, is known as low photon coherence which is a valid approximation for nighttime operation. For convenience, the dark current in this study will be the combined effects due to the detector dark current and due to background radiation. Finally, the thermal or Johnson noise current $i_{th}(t,r)$ is caused by thermal fluctuations of electrons in a resistor. It is modelled as a white Gaussian random process (Ref. 5:145).

To describe the output current in Eq. (5) statistically is very difficult. It is difficult because the output current is a mixed process (i.e., a sum of a discrete process (Point-Process model) and a continuous process (Gaussian model)). Hence, limiting cases must be considered. This study will ignore the current due to thermal noise since one of the objectives of this study is to observe the quantum effects of dark current in filter performance. Also, if the detector is a photomultiplier or avalanche photodiode with high gain, one can assume for low input intensity that the optical receiver is operating in the point process regime, and thus, thermal noise is ignored.

Photodetector Statistics. It is assumed in the detector model that the electrons arriving at the anode of each detector of an array are independent. Then the output statistics of each detector can be treated independently of one another. One needs the statistics describing the number of events, n, the ordered event times, t_k's, and, ordered event locations,

r_k's, to characterize the detector output completely. Thus, the joint density of these variables is needed. With the results given in Eqs. (1) through (3), the above joint density of the photodetector can be derived. Before stating the results, it is convenient to define a complex envelope which characterizes (in part) the focussed or received field (i.e., a classical description). Let $u(t,r)$ denote the scalar optical field. Then the complex envelope $U(t,r)$ of the scalar field is defined implicitly as

$$u(t,r) = \text{Re}\left(U(t,r)e^{j2\pi f_o t}\right) \qquad (6)$$

where $R(\cdot)$ denotes the "real part of" the quantity in parenthesis, and f_0 denotes the carrier frequency. The intensity of the field is represented as

$$I(t,r) = \frac{1}{2} \ \frac{1}{Z_m} |U(t,r)|^2 \qquad (7)$$

where $|\cdot|$ denotes the modulus of the complex field, and Z_m denotes the impedance of the medium (Ref. 2:10).

With the definition of a complex envelope and using Eqs. (6) and (7), the output of the photodetector, conditioned on knowing the optical field, can be modelled as a point process (Ref. 2) with rate function (in one dimension)

$$\lambda(t,r) = \underbrace{\frac{\eta}{hf_0} \int_\ell |U(t,r)|^2 d\ell}_{\text{due to signal}} + \underbrace{\lambda_d}_{\substack{\text{due to dark current} \\ (\lambda_d(t,r) = \lambda_d)}} \qquad (8)$$

20

where n is quantum efficiency, h is Planck's constant, f_o is the optical signal frequency, $U_s(t,r)$ is the complex envelope of the scalar field incident on the detector surface at (t,r), ℓ is the length of detector array, and $d\ell$ is the differential detector length. The integral in Eq. (8) represents the total optical power deposited by the beam on the detector surface and the energy of a photon is hf_o. Also, recall from the previous section that the homogeneous dark rate function is due to the combined effects of background radiation (of zero mean electric field) satisfying low photon coherence and dark current resulting from imperfection of the photodetector. This dark rate function, denoted as λ_d, which is assumed to be independent in time and space can be expressed as

$$\lambda_d = \frac{\eta}{hf_o} \int_\ell E\left[|U_b(t,r)|^2\right]d\ell + \lambda_d' \qquad (9)$$

$$\underbrace{\text{due to background}}_{\text{radiation}} \qquad \underbrace{\text{due to}}_{\text{photodetector}}$$

Note that in Eq. (8) and (9) the optical receiver assumes to be operating at the point-process regime. Also, the first term at the right-hand side of Eq. (9) is a constant rate function, since this study assumes a constant background intensity. In other words, the dark current, λ_d, is independent of time and space. Finally with Eqs. (1) through (3), the output statistics of the optical receiver conditioned on knowing the input field is found in Appendix A as

21

$$f[\{t_k\},\{r_k\}, n] = \exp(-\int_\ell \int_0^T \lambda(\alpha,\beta)d\alpha d\beta) \prod_{k=1}^{n} \lambda(t_k,r_k) \quad (10)$$

where $\{t_k\}$ denotes set of ordered event times, $\{r_k\}$ denotes the corresponding set of event locations, ℓ is detector length, n is the number of events in a time interval $[0,T]$ on ℓ. Again, this joint density completely characterizes the statistical output of the optical receiver.

The Snyder Filter

The previous sections described the statistical model of the optical receiver which is used in the optical tracking problem. The lens system of a receiver focusses the received field on a detector surface at the focal plane. The array of detectors responds to the received field which arrives at various angles at the receiver. Because of random disturbances, the spot wanders randomly on the detector plane. Electrons are released at a rate proportional to the light intensity, given by Eq. (8). For the model upon which the Snyder Filter is based, λ_d is assumed to equal zero. The Snyder Filter estimates the position of maximum intensity on the detector plane in terms of these time-space photoelectrons. An outline of its models and results will now follow specializing for one spatial dimension.

Characterization of Spot. The spot of light on the detector surface is a diffraction pattern focussed by the receiver's lens. It is assumed that the spot of light has a Gaussian-intensity profile given by

$$I(r,\underset{\sim}{x}(t),t) = I_0(t) \ \exp[-\frac{1}{2}(\frac{r-x(t)}{2\,\sigma(t)})^2] \tag{11}$$

where $I_0(t)$ is the maximum intensity, $\sigma(t)$ characterizes the beam spread, and $\underset{\sim}{x}(t)$ is the true location of the spot (i.e., the location at the centroid of the beam). For this study, $I_0(t)$ and $\sigma(t)$ are assumed to be static (i.e., $I_0(t) = I_0$ and $\sigma(t) = \sigma$). Figure 4 portrays this intensity profile. From the figure, the beam width (or beam radius) in this study is defined as a point from the maximum intensity, I_0, to a point in which the intensity is reduced to $I_0 e^{-1/2}$, or about 60.7% of the maximum intensity.

Characterization of Spot's Motion. The motion of the spot's centroid is assumed to be modelled by the following linear stochastic differential equation

$$d\underset{\sim}{x}(t) = F(t)\underset{\sim}{x}(t) + G(t)d\underset{\sim}{\beta}(t); \ \underset{\sim}{x}(t_o) = x_0 \tag{12}$$

where $\underset{\sim}{\beta}(t)$, $t \geq 0$ is a standard Wiener process, $F(t)$ is the state coefficient, and $G(t)$ is the noise input coefficient. The spot's dynamics are governed by a number of disturbances. The atmosphere is one such disturbance; characterized as an inhomogeneous random optical channel (Ref. 9:16-26), the atmosphere changes its index of refraction due to temperature variation. Because of the stochastic nature of the atmosphere, and frequency at which the field propagates, the optical wave interacts with the foreign particles in the atmosphere. One consequence of this phenomenon is that the received field causes variations in the beam's arrival angle, causing the spot to move randomly on the detector plane. Platform
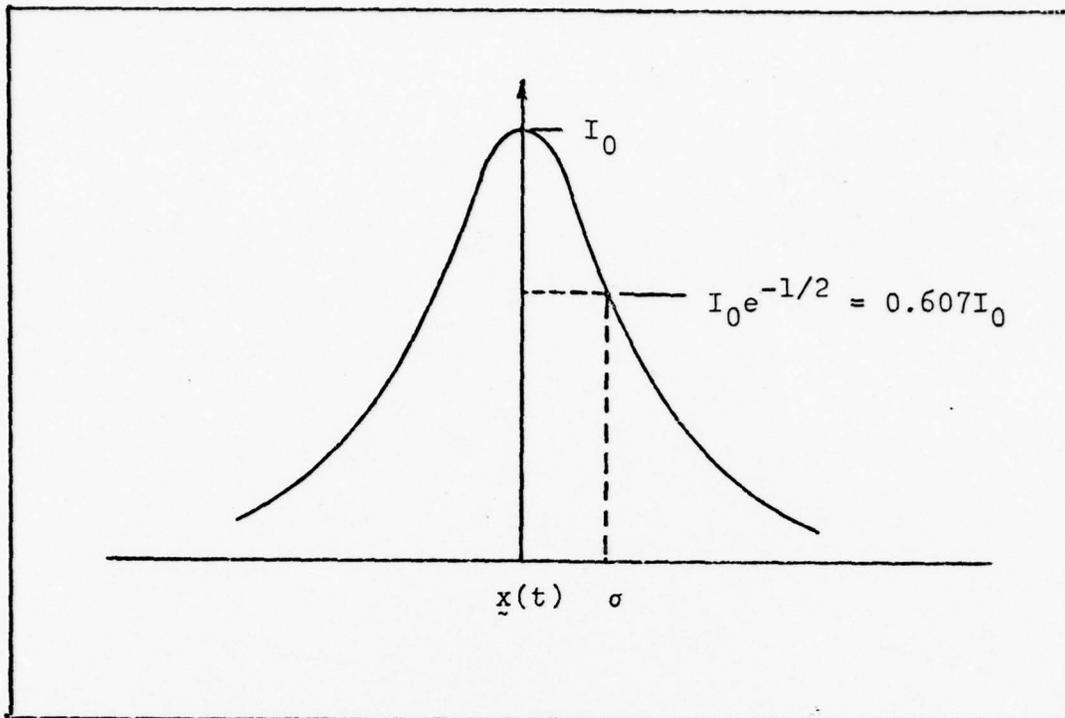
23

Figure 4.   Intensity Profile of Spot
in One-Dimension

vibration of the optical receiver is another disturbance which

causes the spot to drift. The important point is that any

relative motion between object and receiver due to other

effects will result in the spot to drift dynamically on the

detector plane. However, it is not the purpose of this study

to model accurately the motion of the spot, but to provide a

reasonable and simple model for digital simulation. A model

was developed to describe the dynamics of the target relative

to the receiver (Ref. 10:42). This model is a stationary

First Order Gauss-Markov process, the output of a first order

lag driven by white Gaussian noise of zero mean. This process

also describes a variety of phenomena having bandlimited

characteristics. Thus, this model will be used to describe the

dynamics of the spot. Another name for this process is an exponentially time-correlated Gaussian process. Figure 5 portrays the properties of a First Order Markov Process. Two parameters of interest to describe the output of the first order lag are the coherence time τ_c, and its mean squared value σ_x^2. The coherence time (or correlation time), describes how fast the resulting output process, $x(t)$, will vary in time. The reciprocal of τ_c is viewed as the bandwidth of the disturbance process (i.e., turbulance plus other effects). Heuristically, the spot's rms value of random displacements (or spot's rms jitter), is a measure of how much the spot will fluctuate about its mean position on the detector array. Finally, the strength of a white Gaussian noise driving the first order lag is given as

$$E[w(t)w(t+\tau)] = Q\delta(\tau) \tag{13}$$

where $Q = \frac{2}{\tau_c}\sigma_x^2$ (Ref. 9:4.77). The strength, Q, describes how fast the Wiener Process or Brownian Motion, $\beta(t)$, diverges in mean square value from its initial condition of zero (Ref. 9:4.29).

Position-Sensitive Measurements. As mentioned earlier in the report, these measurements consist of event locations which provide information about the position of maximum intensity. Assuming no dark current, the conversion rate at the detector surface is modelled as
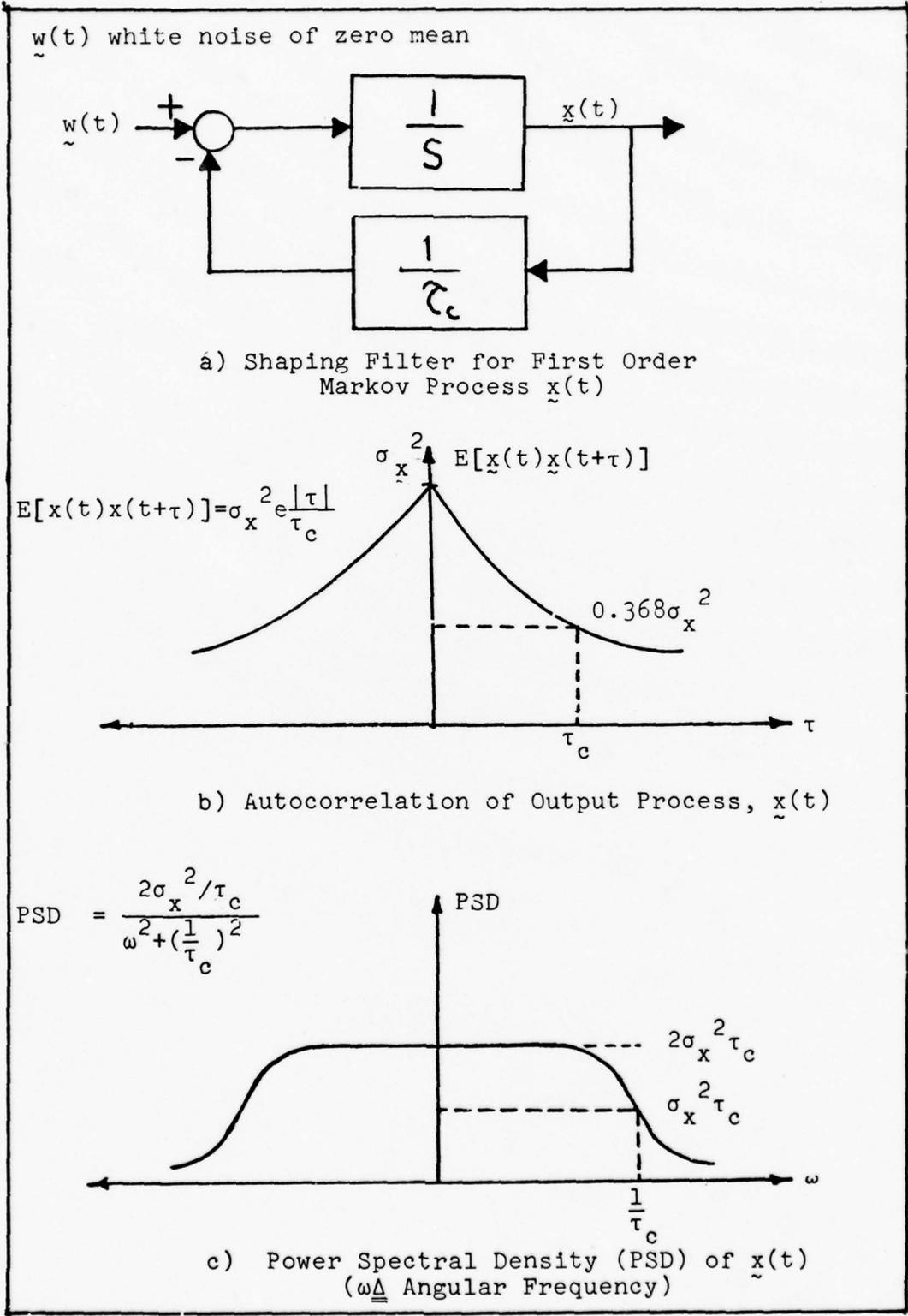
w(t) white noise of zero mean

a) Shaping Filter for First Order
Markov Process $\underset{\sim}{x}(t)$

$$E[x(t)x(t+\tau)]=\sigma_x^2 e^{\frac{|\tau|}{\tau_c}}$$

b) Autocorrelation of Output Process, $\underset{\sim}{x}(t)$

$$PSD = \frac{2\sigma_x^2/\tau_c}{\omega^2+(\frac{1}{\tau_c})^2}$$

c) Power Spectral Density (PSD) of $\underset{\sim}{x}(t)$
($\omega \underline{\underline{\Delta}}$ Angular Frequency)

Figure 5. Properties of a First Order Markov Process

$$\lambda_g(r, \underset{\sim}{x}(t)) = \left(\frac{\eta I_0 \sqrt{2\pi}\, \sigma}{hf_0} \right) \left[\frac{1}{\sqrt{2\pi}\, \sigma} \exp\{ \frac{-1}{2\sigma^2} (r - \underset{\sim}{x}(t))^2 \} \right] \quad (14)$$

If the detector length is sufficiently large, the quantity in brackets, $[\cdot]$, in Eq. (14) integrates over the entire detector length to a value very close to one. Note that Eq. (14) is derived by substituting Eq. (11) as the signal intensity in the first term of Eq. (8). The term $\eta I_0 \sqrt{2\pi}\, \sigma$, represents the average optical energy deposited on the detector by the spot to release photoelectrons. The term hf_0 represents the energy of a single photon, characterizing the particle nature of light of frequency f_0. Then integrating Eq. (14) both over time (with limits from 0 to T) and space (over entire detector length), yields $(\frac{\eta I_0}{hf_0} \sqrt{2\pi}\, \sigma)T$. This term represents the ensemble average of the number of photons detected at the output over the time-space observation area. Recall that the data measurements are obtained from a photo-detector array, capable of measuring the exact temporal and spatial component of a photoelectron. With this idealization the estimator's performance, derived in Appendix B, places a limit in performance which can be achieved with arrays of finite resolution. In other words, a decrease in spatial resolution should result in poorer performance. It is also assumed that the detector array is large such that edge effects of the spot's intensity profile, described as Gaussian-Shaped can be ignored. A good "rule of thumb" is to have the detector large such that the distance between the center of the spot

27

and the edge of the detector array is at least six beam
widths. This is "good" because with respect to a Gaussian-
shaped beam, the above "rule of thumb" closely approximates
a detector of infinite length. Also, by having the detector
length large, edge effects may be ignored as the spot moves
dynamically even with large wanderings on the detector surface.

Snyder's Filter Equations. This subsection presents the
dynamical filter equations that incorporate point-process
observations to estimate the position of maximum intensity.
These equations, specialized for one-dimension state as follows:

$$d\hat{x}(t) = F(t)\hat{x}(t)dt + \int_\ell K(t)[r-H(t)\hat{x}(t) \, N(dtxd\ell) \quad (15)$$

$$d\hat{P}(t) = F(t)\hat{P}(t)dt + \hat{P}(t)F(t)dt + G(t)Q(t)G(t)dt$$

$$- \int_\ell K(t)H(t)N(dtxd\ell) \quad (16)$$

$$K(t) = \hat{P}(t)H(t)[H(t)\hat{P}(t)H(t) + R(t)]^{-1} \quad (17)$$

$$\hat{x}(t_0) = 0, \, P(t_0) = P_0 \quad \text{(initial conditions at } t_0) (18)$$

where $\hat{x}(t)$ is the conditional estimate of the spot's centroid
given the measurement data (i.e., event locations that are
random in time and space), $\hat{P}(t)$ is the conditional error
variance, $H(t)$ is the measurement coefficient, $K(t)$ is the
filter gain, r is the location of the photoconversion from a
differential detector length $d\ell$ in the differential time
dt, and $N(dtxd\ell) = dN$ is the differential number of events
in the differential "space-time" observation area, $dtxd\ell$.

28

The filter equations, Eq. (15) through (18) are recursive relations, similar in form to the discrete-time measurement Kalman-Bucy Filter. They differ in that the Snyder Filter updates its estimate and uncertainty at the event times whereas the Kalman-Bucy Filter updates at predetermined times. The mean and the variance of the conditional probability density function which describes the error in the estimate of the spot's centroid are propagated forward from the time period of the kth photoconversion t_k^+, to the time of the $(k+1)^-$st photoconversion, t_{k+1}^-. Then the filter updates at time t_{k+1}, by processing the measurement at that time, to achieve the estimate $\hat{x}(t_{k+1}^+)$, where "−" denotes before measurement update and "+" denotes after measurement update

It is convenient to expand further the interpretation of the filter equations. During interarrival times between photoconversions, the filter's conditional estimate of the centroid's position $\hat{x}(t)$, and its error variance $\hat{P}(t)$, evolve in time by the following equations

$$\frac{d\hat{x}(t)}{dt} = F(t)\hat{x}(t) \qquad\qquad ; \ t_k < t \le t_{k+1} \qquad (19)$$

$$\frac{d\hat{P}}{dt} = 2F(t)\hat{P}(t) + G(t)Q(t)G(t) \qquad ; \ t_k < t \le t_{k+1} \qquad (20)$$

The above relations govern how $\hat{x}(t)$ and its associated conditional error variance $\hat{P}(t)$ will propagate when the filter does not observe a photoconversion. But when the filter does observe an event, via the detector array, it will update $\hat{x}(t)$ and $\hat{P}(t)$ after the occurrence of the event by the following update

equations:

$$\hat{P}(t_k^+) = [H(t_k)R^{-1}(t_k)H(t_k) + \hat{P}(t_k^-)]^{-1} \qquad (21)$$

$$\hat{x}(t_k^+) = \hat{P}(t_k)[H(t_k)R^{-1}(t_k)r_k + \hat{P}^{-1}(t_k^-)\hat{x}(t_k^-)] \quad (22)$$

where t_k^+ is the time after the occurrence of the kth photoconversion, t_k^- is the time before this occurrence, and r_k is the location of photoconversion (or photoelectron). By using the matrix inversion lemma (Ref. 1), Eqs (21) and (22) may be alternatively expressed as

$$\hat{x}(t_k^+) = \hat{x}(t_k^-) + K(t_k^-)[r_k - H(t_k)\hat{x}(t_k^-)] \qquad (23)$$

$$\hat{P}(t_k^+) = \hat{P}(t_k^-) - K(t_k^-)H(t_k)\hat{P}(t_k^-) \qquad (24)$$

where

$$K(t_k) = \hat{P}(t_k^-)H(t_k)[H(t_k)\hat{P}(t_k^-)H(t_k) + R(t_k)]^{-1} \quad (25)$$

Eqs. (22) and (23) indicate, clearly, the recursive nature of how $\hat{x}(t)$ and $\hat{P}(t)$ evolve in time given the history of data measurements. In other words, the "recursive" nature of the Snyder Filter does not require all previous data measurements to be kept in storage and processed every time a photoconversion is observed. This saves computer resources which is an important consideration for filter implementaton.

To summarize, the Snyder Filter is a recursive data processing algorithm, using the available information from the photoconversions to improve the accuracy of estimating

the spot's current position of maximum intensity. The Snyder
Filter is a minimum mean square error estimator; it attempts
to provide the best estimate possible in such a manner that
the error is minimized statistically.

The next section specifies parameters where sufficient
conditions are presented to provide actual mean-square-error
(MSE).

Upper and Lower Bounds

A question should be asked whether or not the filter
errors are small enough for tracking purposes. Thus, the
MSE (i.e., an expectation of the conditional error variance
given the data measurements over all possible histories of
measurements taken), is used to make this evaluation.
Unfortunately, finding the mean-square-error cannot be evaluated
in closed form. Rather than attempt to solve for the MSE,
one can bound it above and below. The method involves easily
precomputed upper and lower bounds ($P_* \triangleq$ lower bound on MSE and
$P^* \triangleq$ upper bound on MSE).

The following is a quote from Ref. (4) in how these
bounds are derived:

> "The upper bounds are derived by evaluating
> exactly the performance of a parameterized family
> of suboptimum design; one of these is identified
> as having smaller performance than any other, thus
> providing a minimal upper bound within this family
> of suboptimum design. The lower bounds are obtained
> directly by calculations involving inequalities."

The above results are specialized for this open-loop problem
in one dimension. Here, the upper and lower bounds are

solved in steady state (i.e., $\frac{dP_*}{dt} = \frac{dP^*}{dt} = 0$) and are
derived in Appendix B. The solution to this steady-state
estimation problem involves a quadratic equation yielding a
result given as

$$\frac{P_*}{\sigma^2} = \frac{1}{2X} \left[-1 + (1 + 4XY)^{1/2} \right] \tag{26}$$

$$\frac{P^*}{\sigma^2} = \frac{1}{2(X+1)} \left[(Y-1) + \{(Y-1)^2 + 4(X+1)Y\}^{1/2} \right] \tag{27}$$

where

$$X = (\frac{\eta}{hf_0} I_0 \sqrt{2\pi} \sigma)(1/2\tau_c) \tag{28}$$

$$Y = \frac{\sigma_x^2}{\sigma^2} \tag{29}$$

Several comments deserve mentioning with regard to Eqs.
(26) through (29). The upper and lower bounds are normalized
with respect to the square of the beam width. This normalization
is convenient since the error in estimating the spot can
be measured in terms of beam width (i.e., a performance or
criterion can be established). Thus, for good tracking
performance, it is desirable to have $\frac{P_*(t)}{\sigma^2}$ and $\frac{P^*(t)}{\sigma^2}$ small.
Figures (6a-6i) are plots of the upper and lower bounds on
MSE as a function of X and Y. Note the absolute magnitude
and relative separation of the bounds as X and Y are varied
on each of these plots. Consider Figs. 6a-6d : for a fixed
Y, the bounds both decrease in magnitude and in relative

32
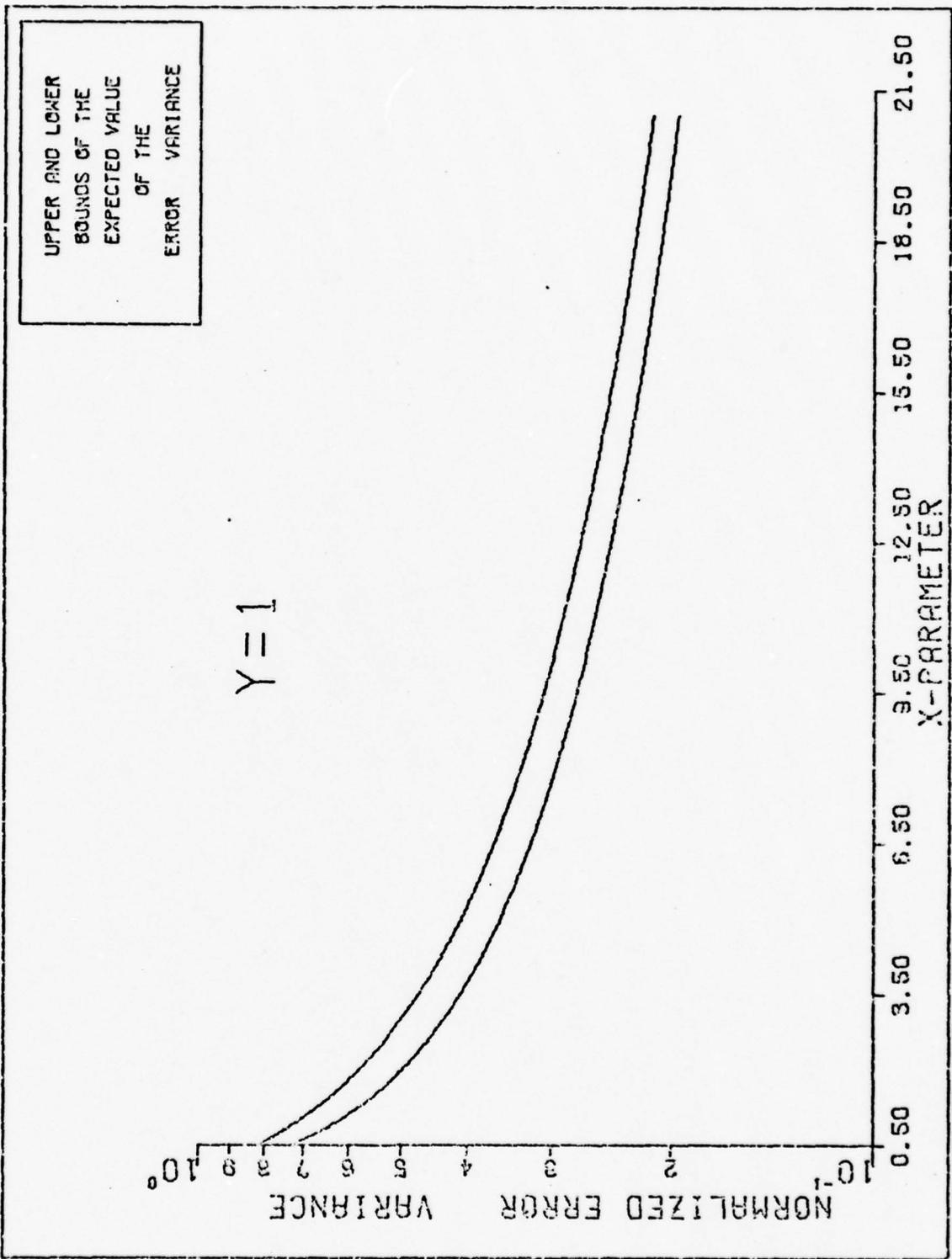
Figure 6a. Upper and Lower Bounds on MSE

Y=5

X-PARAMETER

NORMALIZED ERROR VARIANCE
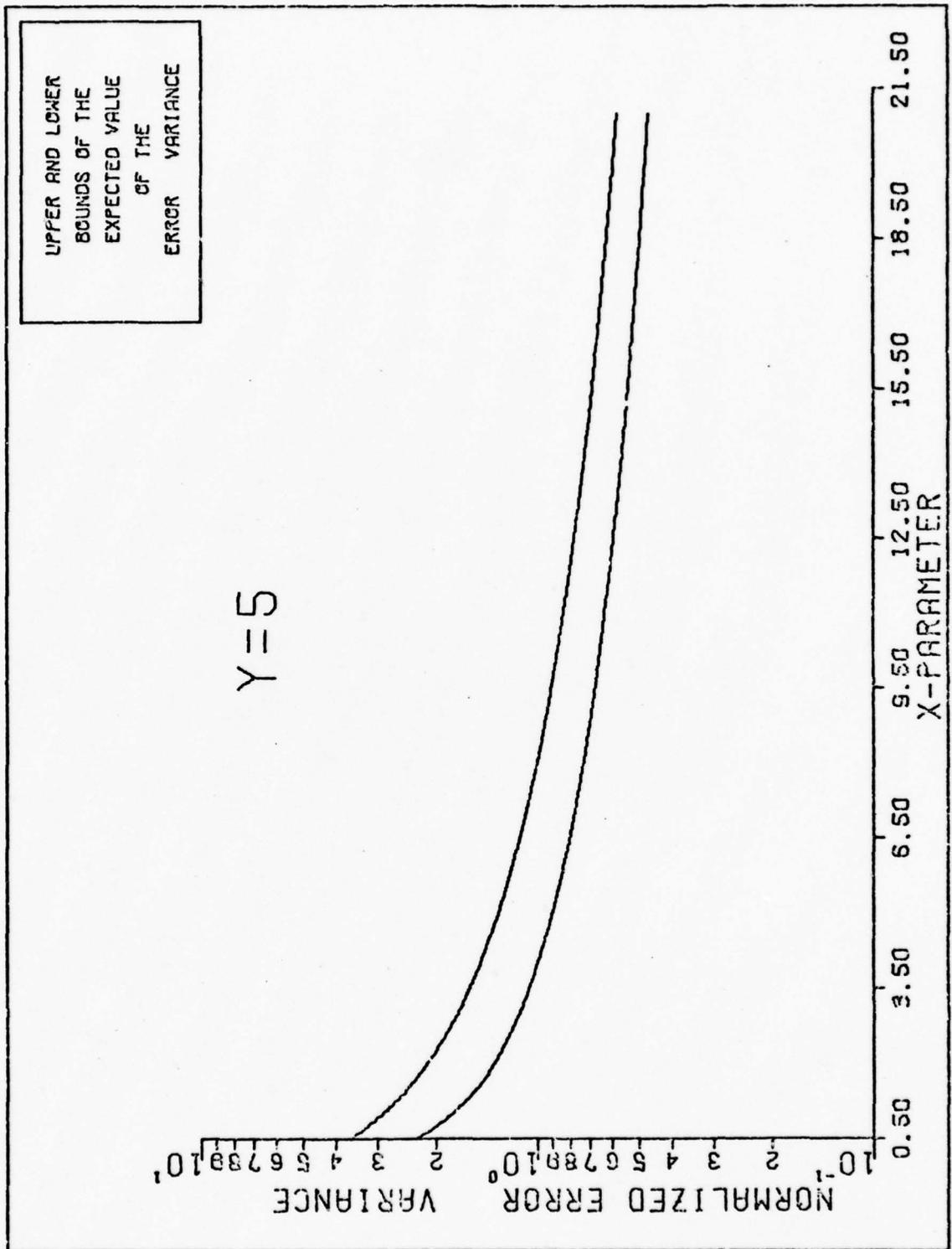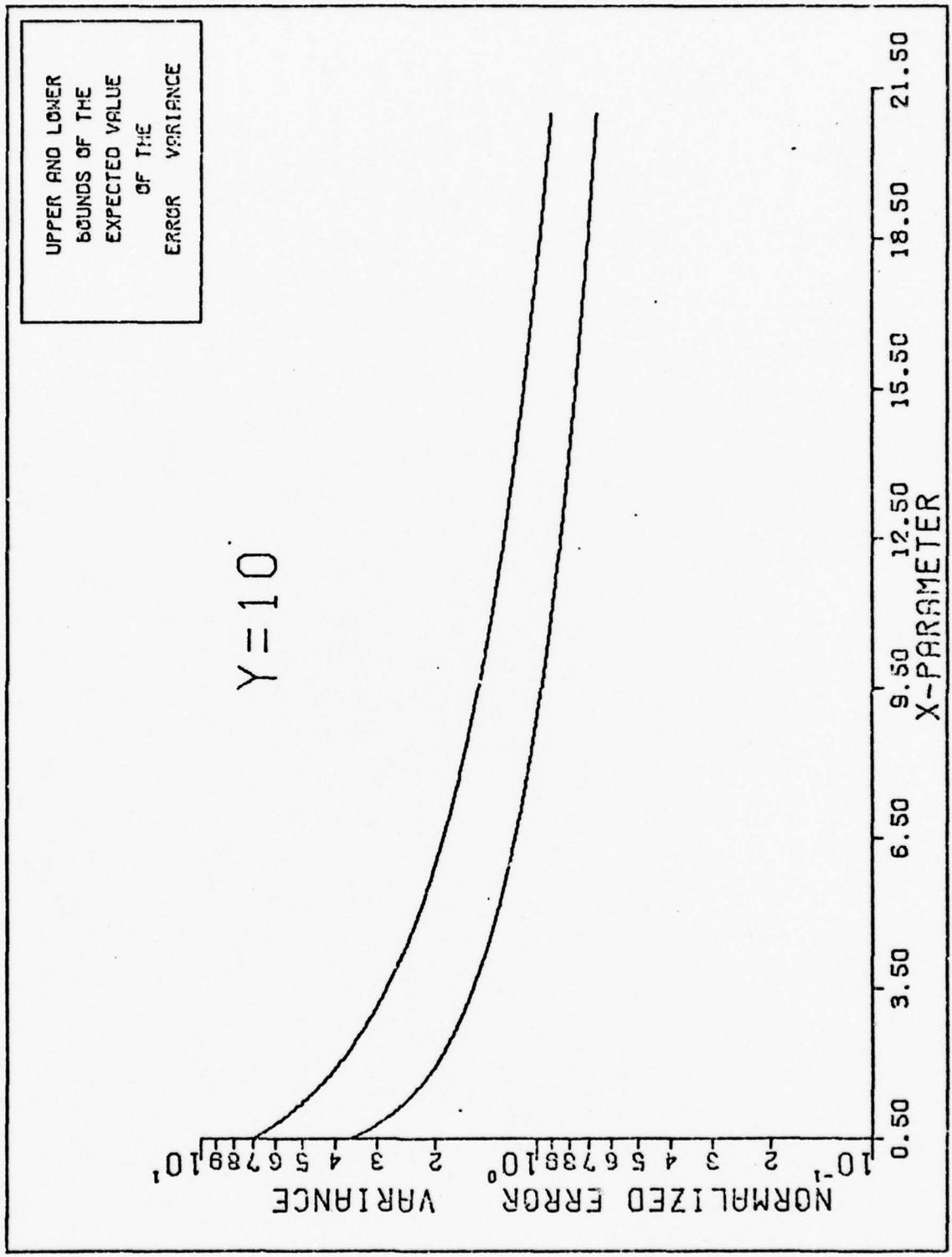
Figure 6b. Upper and Lower Bounds on MSE

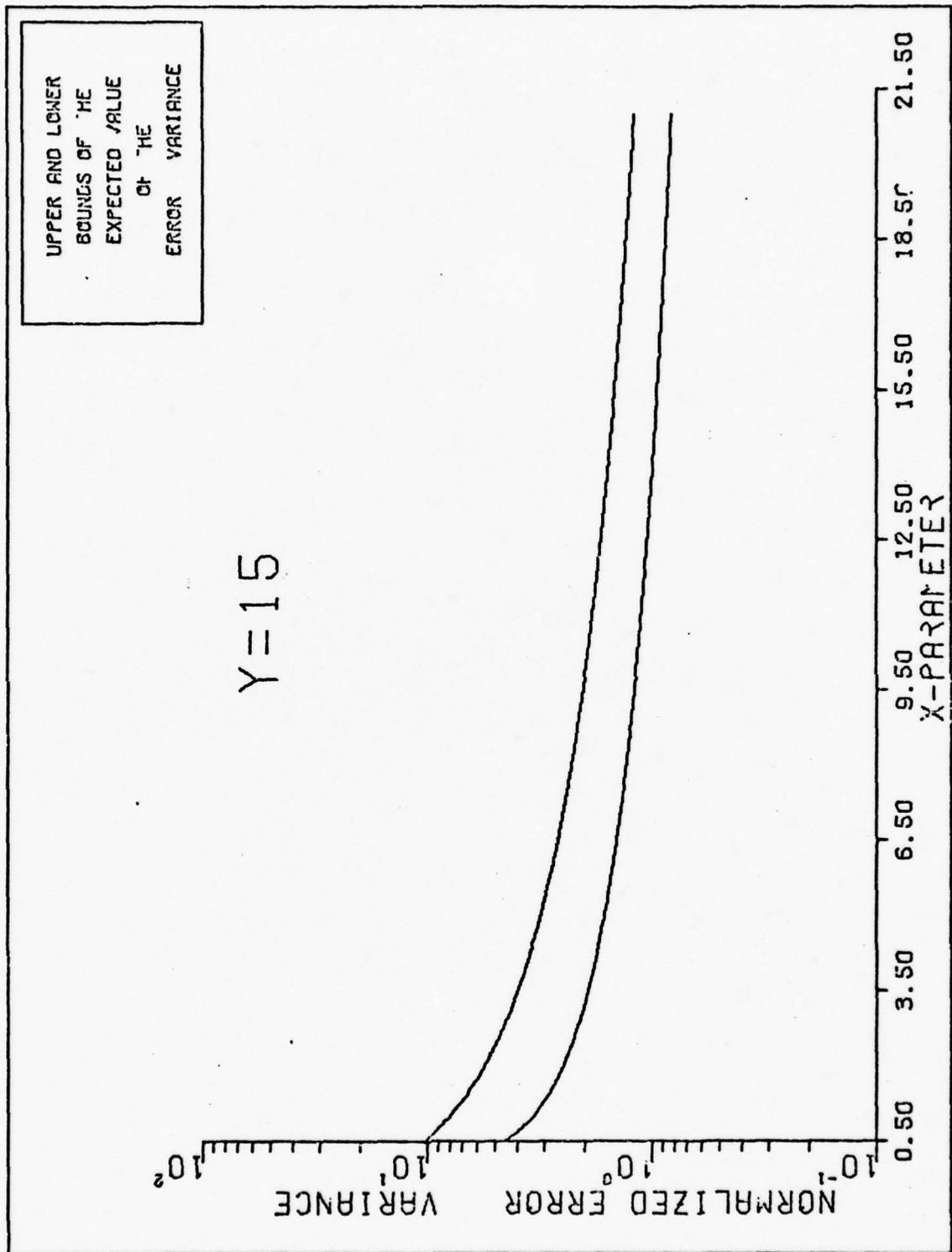Figure 6c. Upper and Lower Bounds on MSE

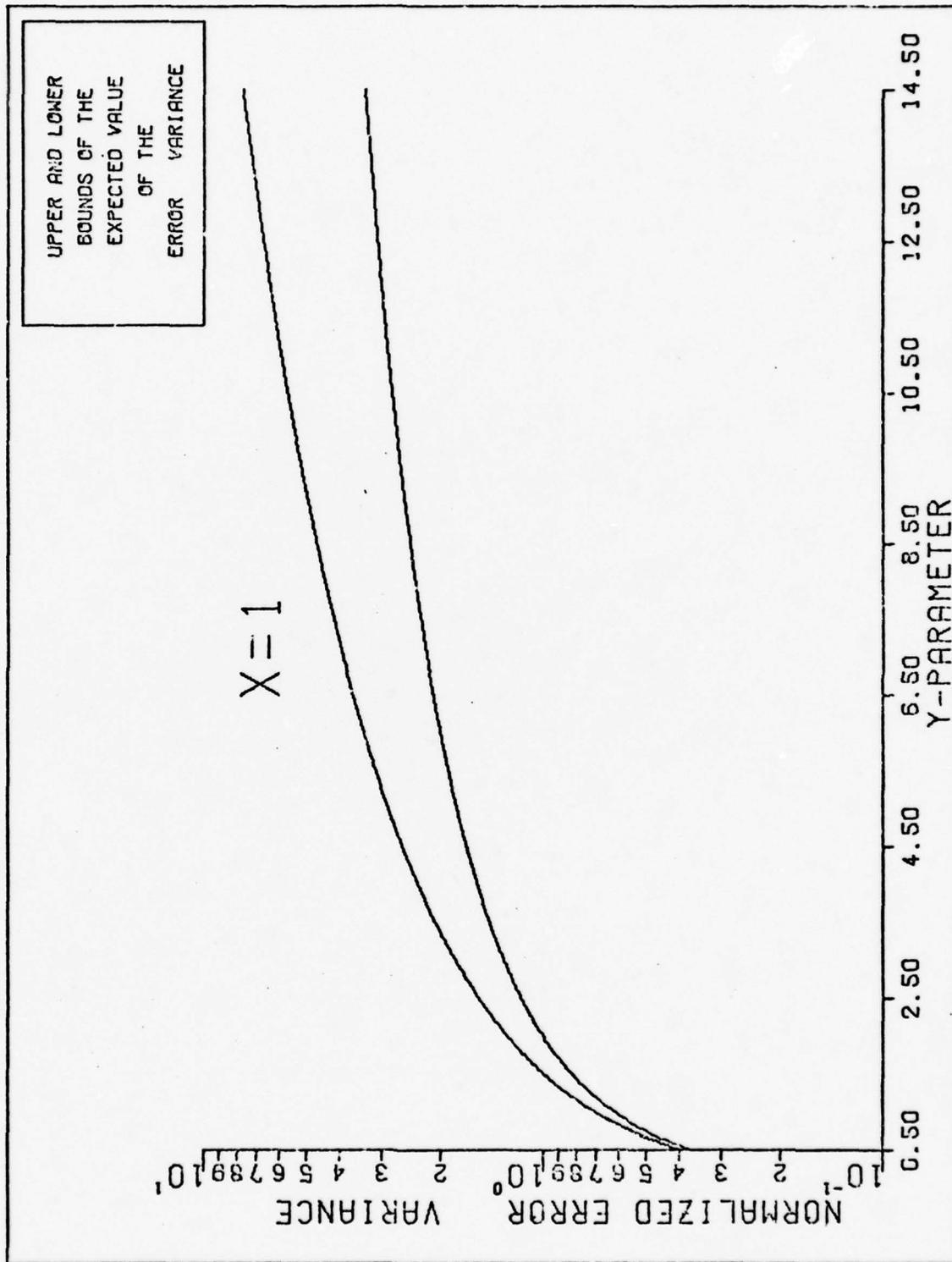Figure 6d.   Upper and Lower Bounds on MSE
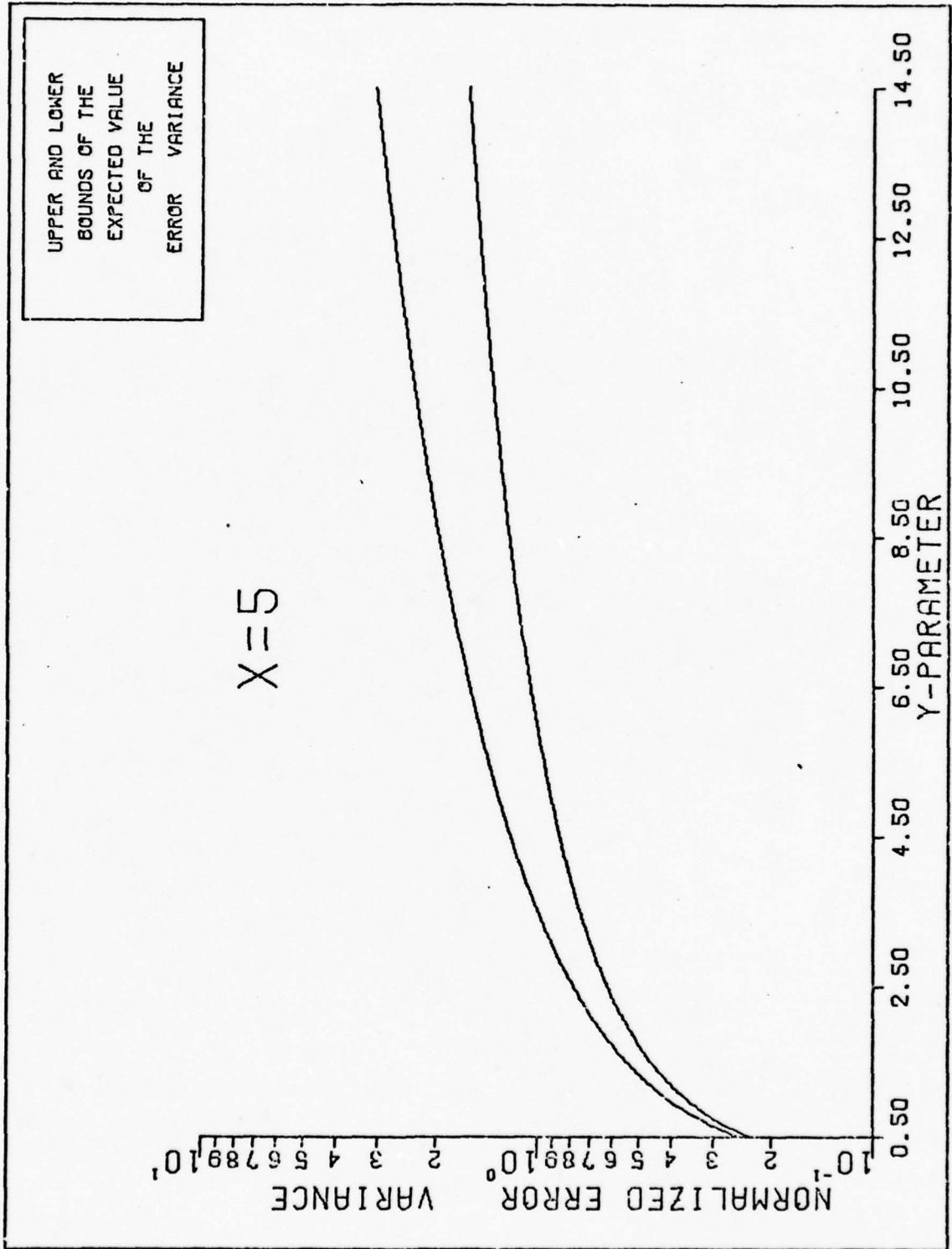
Figure 6e. Upper and Lower Bounds on MSE
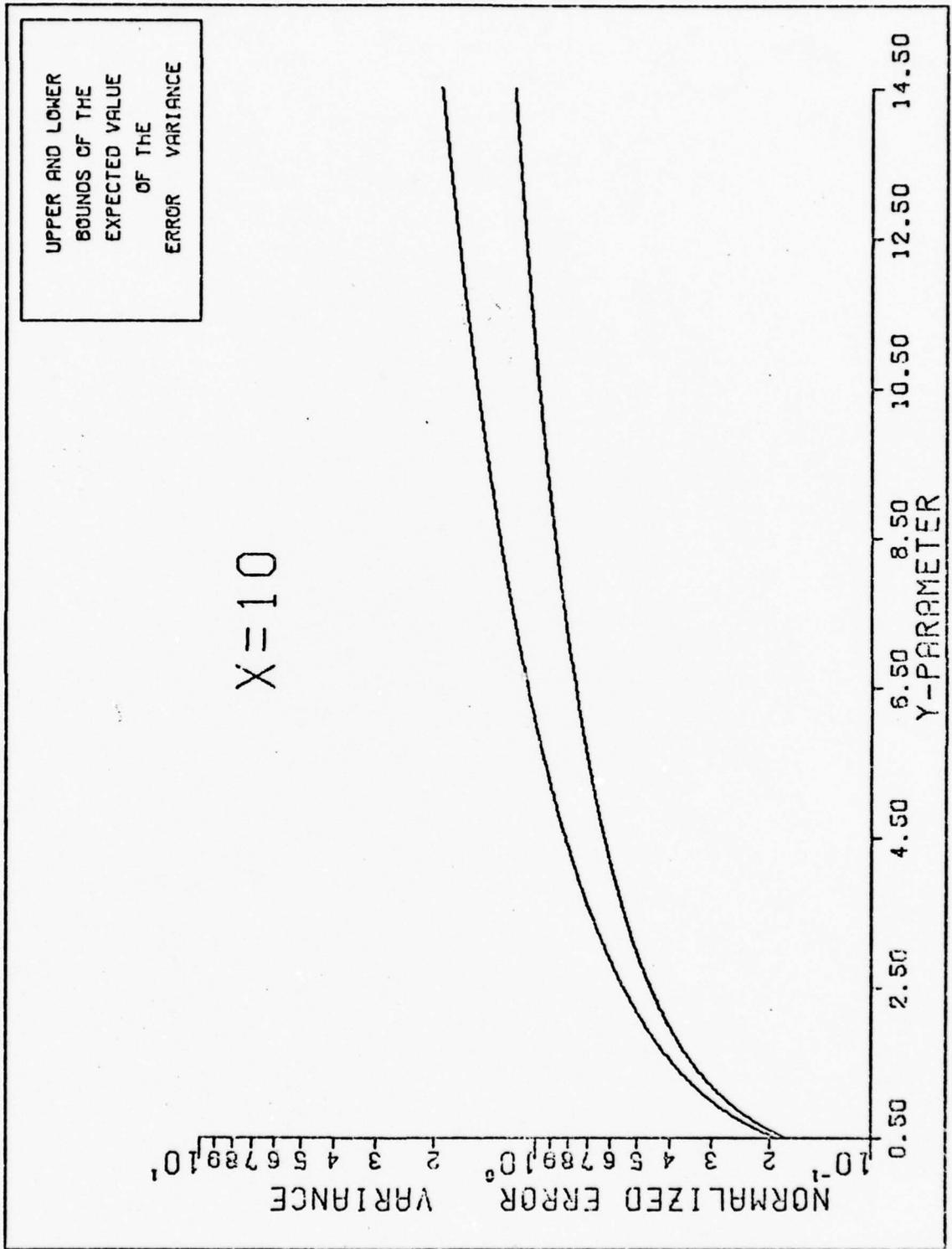
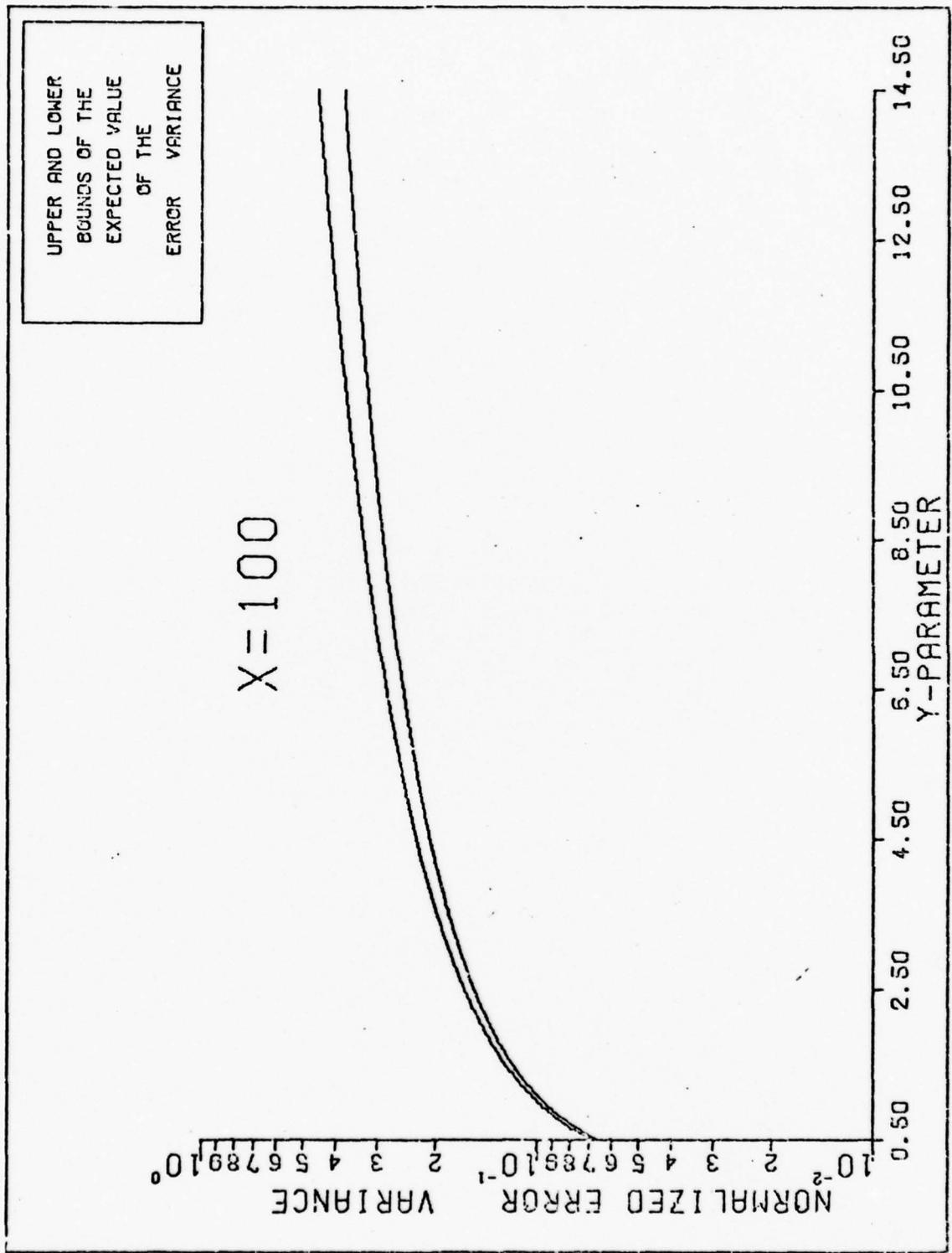Figure 6f. Upper and Lower Bounds on MSE

Figure 6g.   Upper and Lower Bounds on MSE
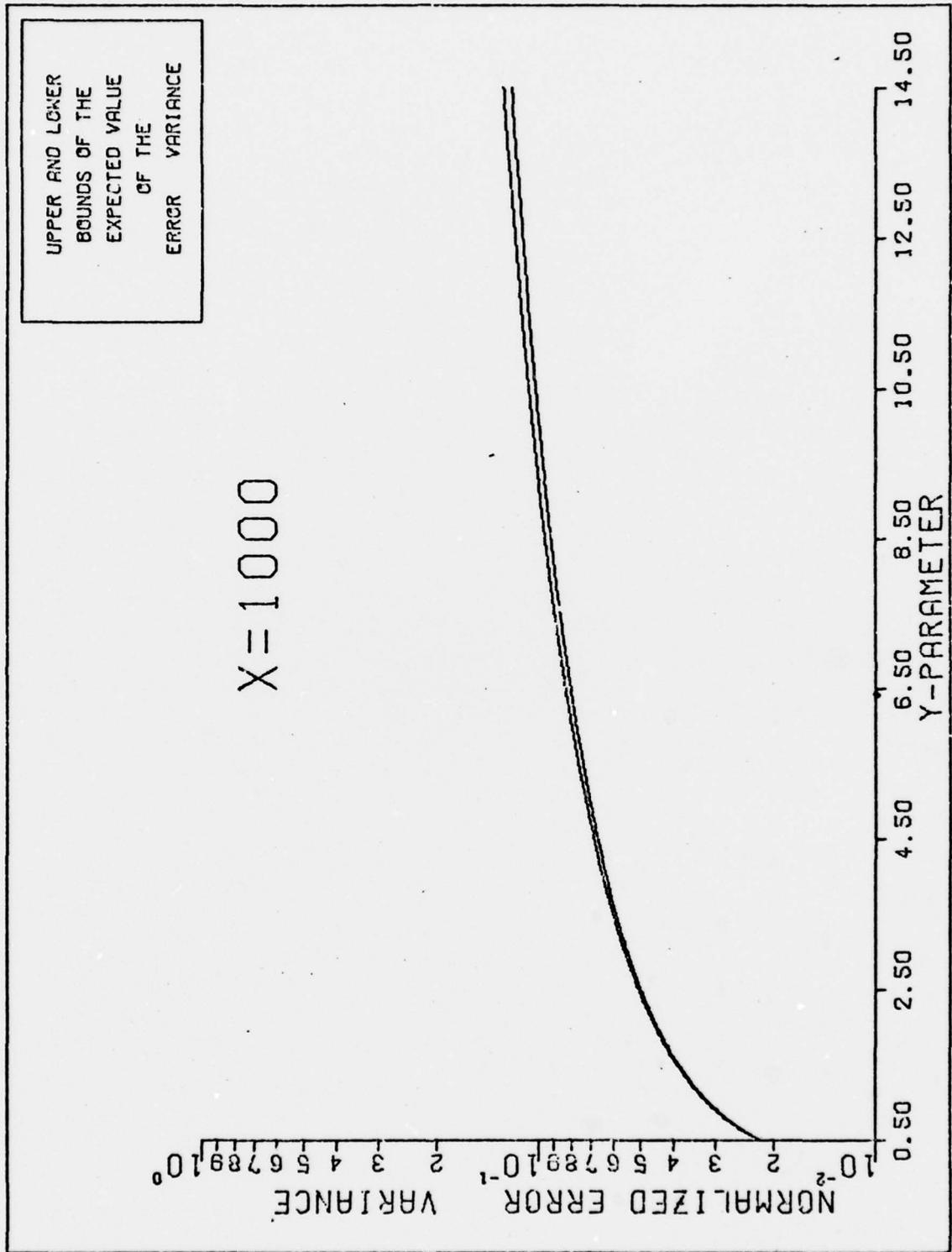
Figure 6h. Upper and Lower Bounds on MSE

Figure 61.   Upper and Lower Bounds on MSE

seperation as X gets larger. Physically, the X-parameter can be described as the number of photoconversions in a coherence time of the spot's dynamics. As the number of data observations increases (i.e., X-parameter increases), the filter has more information to correct and update its estimate of the spot's current position of maximum intensity. Thus, the filter can update its estimate frequently if a large number of events are observed in a relatively short period of time. Suppose, for example, the spot drifts on the detector with large "wanderings". Also, assume the filter has just processed one event and does not observe another event in a relatively long time (i.e., a low rate function or a small X). Then the information provided by the "processed" event, has little value at some later time if the spot is "jittering" at higher frequency with large "wanderings" (i.e., spot dynamics have a small coherence time, τ_c, and a large jitter, σ_x^2). On the other hand, if the filter frequently observes a larger number of events within a coherence time, the Snyder Filter has more information concerning the spot's location on the detector surface.

Consider now a fixed X and a variation of the Y-parameter (Fig. 6e-6i). As Y increases, the bounds increase in magnitude and in relative separation. Y is the ratio between the spot's jitter normalized by the square of the beam width. The beam width indicates whether the filter should track the event locations closely in order to locate the spot. As the beam

42

width increases, photoelectrons are distributed at larger distances from the spot's centroid. Hence, increasing the beam width results in having less weight in the point-process measurements. Thus, the optimal gain, $K(t)$, should decrease as seen in Eq. (17).

Similarly, the noise strength, σ_x^2, indicates stronger noise sources driving the dynamics of the spot or increasing the uncertainty in the model. Increasing σ_x^2 increases both the rate of growth of the conditional error covariance, $\hat{P}(t)$, between event times. This results in increasing values of the filter gain, implying larger weighting of the point process measurements. Stated another way, increasing the jitter indicates that the filter must put less "confidence" in the prediction capability of the filter's own internal dynamics model.

Physically, it is desirable to have a narrow beam width such that electrons can be distributed near the position of maximum intensity. At the same time, the spot's jitter should be small so that photoelectrons are released from a small area on the detector. This physical reasoning is described by the Y parameter.

With regard to the above discussion and by careful observation of Eqs (26) and (27), the upper and lower bounds converge to the steady state value of the actual MSE when $X \gg 1$ and $Y \ll 1$. Referring back to the performance parameters, the estimator will perform well if a large number of signal photoelectrons are observed frequently and they appear

in a small area of a photodetector surface. This physical
reasoning is described by the above parameters derived from
upper and lower bounds. These parameters establish regions of
operation where the upper and lower bounds converge to the actual
mean square error. However, when the bounds do not converge,
actual performance is difficult to obtain. To demonstrate
the usefulness of the bounds and to determine which bound lies
closer to actual performance, a Monte Carlo simulation is
conducted. These parameters and the Snyder Filter assumes
no dark current in the measurement process; and so an additional
objective is to include dark current in the simulation to
determine its effects in filter performance.

Summary

The chapter began by describing the direct detection optical
receiver. The important point to remember is that the statistics
of the output of the receiver govern the performance of optical
components performing postdetection processing and vice versa.
These statistics are described by a mathematical model called
the point process. The model is applied in situations when the
receiver is operating under low signal and background radiation.
In this case, the discrete nature of light at optical frequencies
and the nature of the detection process fundamentally limit the
performance of optical trackers.

The Snyder Filter was discussed next. This filter, a recursive
data processing algorithm, processes the point process observations
of the receiver to estimate optimally the position of

maximum intensity on the detector surface. The Snyder Filter
assumes its observations are from an infinitesimally fine
detector array (i.e., the detector array has infinite resolution
to measure the exact location of a photoconversion).

The chapter concluded with the performance parameters
describing the mean-square-error of the Snyder Filter. The
parameters are derived from the upper and lower bound on MSE.
The two parameters are defined to establish conditions under
which the Snyder Filter should yield good tracking. One parameter
is the ensemble average of the number of photons detected in a
coherence time of the spot's dynamics, and the other is the
spot's jitter normalized by the square of the beam width. These
parameters also indicate sufficient conditions, such that the
bounds converge to the actual MSE. The MSE can then be used
to evaluate the tracking capabilities of the Snyder Filter
(i.e., to determine whether the average errors committed
by the Snyder Filter are sufficiently small for tracking purposes).
Again, these performance parameters and the Snyder Filter
assume that the events locations are only due to the spot.

III. Simulation of the Snyder Filter and Performance Analysis

The last chapter addressed the following concepts: (1) a mathematical model of the output process of the optical receiver, (2) a model of the spot's intensity profile and its dynamics, (3) the Snyder Filter which incorporates the measurement data $\{t_k\}$ and $\{r_k\}$ from the output of the receiver to estimate the position of maximum intensity, and (4) performance parameters which establish regions of operations where the bounds converge to actual MSE.

This chapter develops algorithms to implement the above models and to simulate the Snyder Filter in a digital computer. The algorithms that were developed are (1) simulation of spot's dynamics, (2) a noise generator, and (3) algorithms to generate the time-space point process observation (two methods).

Concluding the chapter is the performance analysis used to evaluate the tracking capabilities of the Snyder Filter.

Simulation of Spot's Dynamics

The spot's dynamics on the detector surface   modelled by a linear stochastic differential equation given in Eq. (12), or in less rigorous form,

$$\dot{x}(t) = F(t)x(t) + G(t)w(t) \tag{30}$$

where $w(t)$ is a white noise with zero mean and variance

$$E[w(t)w(t')] = Q(t)\delta(t-t') \qquad (31)$$

where $Q(t)$ is a diffusion strength variable. Consider an equivalent discrete-time model for digital simulation. Here, equivalence means that a discrete time process has the same second order statistics as $\underset{\sim}{x}(t)$, when viewed at the sample times. As shown in Ref. (9:4.63), an equivalent discrete time solution to Eq. (30) is

$$\underset{\sim}{x}(t_{i+1}) = \phi(t_{i+1},t_i)\underset{\sim}{x}(t_i) + \underset{\sim}{w}_d(t_i) \qquad (32)$$

where t_{i+1}, t_i are the discrete points in time, $\phi(t_{i+1},t_i)$ is the state transition matrix (scalar, in this case), $\underset{\sim}{w}_d(t_i)$ is the white discrete-time stochastic process with statistics:

$$E[\underset{\sim}{w}_d(t_i)] = 0$$

$$E[\underset{\sim}{w}_d^2(t_i)] = Q_d(t_i)$$

$$= \int_{t_i}^{t_{i+1}} \phi(t_{i+1},\tau)G(\tau)Q(\tau)G(\tau)\phi(t_{i+1},\tau)d\tau$$

$$\qquad\qquad ; \ t_i = t_j$$

$$E[w_d(t_i)w_d(t_j)] = 0 \qquad\qquad ; \ t_i \neq t_j$$

From Ref. (9:30), the mean $m_x(t)$ and variance $P_{xx}(t)$ of the process $\underset{\sim}{x}(t)$ propagate in time as

$$m_x(t_{i+1}) = \phi(t_{i+1},t_i)m_x(t_i) \qquad (33)$$

47

$$P_{xx}(t_{i+1}) = \phi(t_{i+1}, t_i)P_{xx}(t_i)\phi(t_{i+1}, t_i) + Q_d(t_i) \quad (34)$$

A simple model of the spot's dynamics on the detector is an output process of a first order lag driven by white Gaussian noise of zero mean. This model is represented as

$$\dot{\underline{x}}(t) = -\frac{1}{\tau_c}\underline{x}(t) + \underline{w}(t) \quad (35)$$

where $F(t) = -\frac{1}{\tau_c}$ and $G(t) = 1$. Then, its state transition matrix is

$$\phi(t, \tau) = \phi(t-\tau) = e^{-\frac{1}{\tau_c}(t-\tau)}$$

The model therefore becomes,

$$\underline{x}(t_{i+1}) = e^{-\frac{\Delta t}{\tau_c}}\underline{x}(t_i) + \underline{w}_d(t_i)$$

where Δt is sample period (or observation time increment), and $\underline{w}_d(t_i)$ is a white Gaussian discrete-time process of zero mean and covariance

$$E[\underline{w}_d^2(t_i)] = \int_{t_i}^{t_{i+1}}\phi(t_{i+1}, \tau)G(\tau)Q(\tau)G(\tau)\phi(t_{i+1}, \tau)d\tau$$

$$= \frac{\tau_c}{2}Q[1 - e^{-\frac{2\Delta t}{\tau_c}}]$$

$$= \sigma_x^2[1 - e^{-\frac{2\Delta t}{\tau_c}}]$$

where $Q = \frac{2}{\tau_c}\sigma_x^2$, and $\sigma_x^2 \triangleq$ mean square value of output process $\underline{x}(t)$, (in this report, σ_x^2, represents the "jitter" of spot). Q is found in Ref. 9:4.77 ; and it represents the

48

the strength of the white noise. Eq. (35) is the expression which simulates the spot's dynamics on the detector. The next section discusses how a white Gaussian discrete-time process $w_d(t)$ is generated.

Noise Generators

A uniform random variable distributed between 0 and 1 is supplied by the computer system. This uniform number generator is used to generate realizations of random variables described by desired probability density functions or probability distribution functions. This will be shown in the next section.

However, applying the Central Limit of Statistics, one can generate an essentially Gaussian random variable by adding twelve independent uniform random variables distributed between 0 and 1. Since the mean and variance of a single uniform random variable is 1/2 and 1/12, respectively, the above addition yields a Gaussian random variable with mean 6 and variance 1. To obtain a zero mean and a unit variance simply subtract 6 from the above result. Note that the Central Limit Theorem implies that the resulting random variable w has a probability distribution that approximates a Gaussian distribution (Ref. 11).

Finally, a Gaussian random variable w of desired variance σ_w and mean μ_w is given by the following transformation equation

$$z = \frac{w - \mu_w}{\sigma_w}$$

or

$$w = \sigma_w z + \mu_w$$

where z is a Gaussian random variable with zero mean and unit variance.

The "whiteness" of the Gaussian discrete-time sequence $\underset{\sim}{w}_d(t_i)$, is obtained by independent calls of the above algorithm The resulting set of w's forms the white-Gaussian discrete-time process $\underset{\sim}{w}_d(t_i)$ (i.e., $E[w_d(t_i)w_d(t_j)] = 0$ for all $t_i \neq t_j$). Also, note that the second order statistics of $\underset{\sim}{w}_d(t_i)$ (or set of random variables w) are the second order statistics of the output process $\underset{\sim}{x}(t)$, of the first order lag. The statistics of $\underset{\sim}{w}_d(t_i)$ are

$$\mu_w = 0$$

$$\sigma_w^2 = \sigma_x^2 \left(1 - e^{-\frac{2\Delta t}{\tau_c}}\right)$$

Point Process Generation in Time and Space

In this report, two methods are employed to generate the event locations in time and space. One, the Binomial Method, is based on the incremental definition of the point process, and the other, the Conditional Method, makes use of conditional probability densities describing the output statistics of the receiver. The latter method uses Bayes' Rule. It should be noted that these conditional densities are derived by the incremental definition of a point process.

50

The Binomial rule will be discussed first. Next, a transformation equation used by the Conditional Method is described; then the Conditional Method follows. Concluding this section is a discussion on the theoretical and practical aspects of the two methods of simulation.

Binomial Method. Shown in Fig. (7) is a flowchart of the Binomial algorithm that generates the time-space events. As mentioned earlier, this development is based on the incremental definition of a point process. It is the same development describing the model of the photodetecting surface. For the purpose of clarity, a discussion on the point process incremental definition in one-dimensional space is given as

(1) Event in a space-time observation area, $\Delta\ell\Delta t$, are statistically independent where $\Delta\ell$ is the incremental detector length, Δt is the incremental time observation.

$$P\begin{bmatrix} \text{one event in} \\ \Delta\ell\Delta t \\ \text{located at } (t,r) \end{bmatrix} = \lambda(r,\underset{\sim}{x}(t))\Delta t\Delta\ell + o(\Delta t\Delta\ell) \qquad (36)$$

$$P\begin{bmatrix} \text{zero events in} \\ \Delta\ell\Delta t \\ \text{located at } (t,r) \end{bmatrix} = 1 - \lambda(r,\underset{\sim}{x}(t))\Delta t\Delta\ell + o(\Delta t\Delta\ell) \qquad (37)$$

where $P[\cdot]$ denotes probability of, the "$o\Delta\ell\Delta t$" denotes higher order terms of $\Delta\ell\Delta t$ (i.e., $\underset{\Delta t\Delta\ell\to o}{\text{limit}} (\frac{o(\Delta t\Delta\ell)}{\Delta t\Delta\ell}) = 0$), and $\underset{\sim}{x}(t)$ denotes location of maximum intensity on the detector surface. Note that different values of r will give different values of rate functions, $\lambda(r,\underset{\sim}{x}(t))$.

Consider a uniform random number distributed between zero and
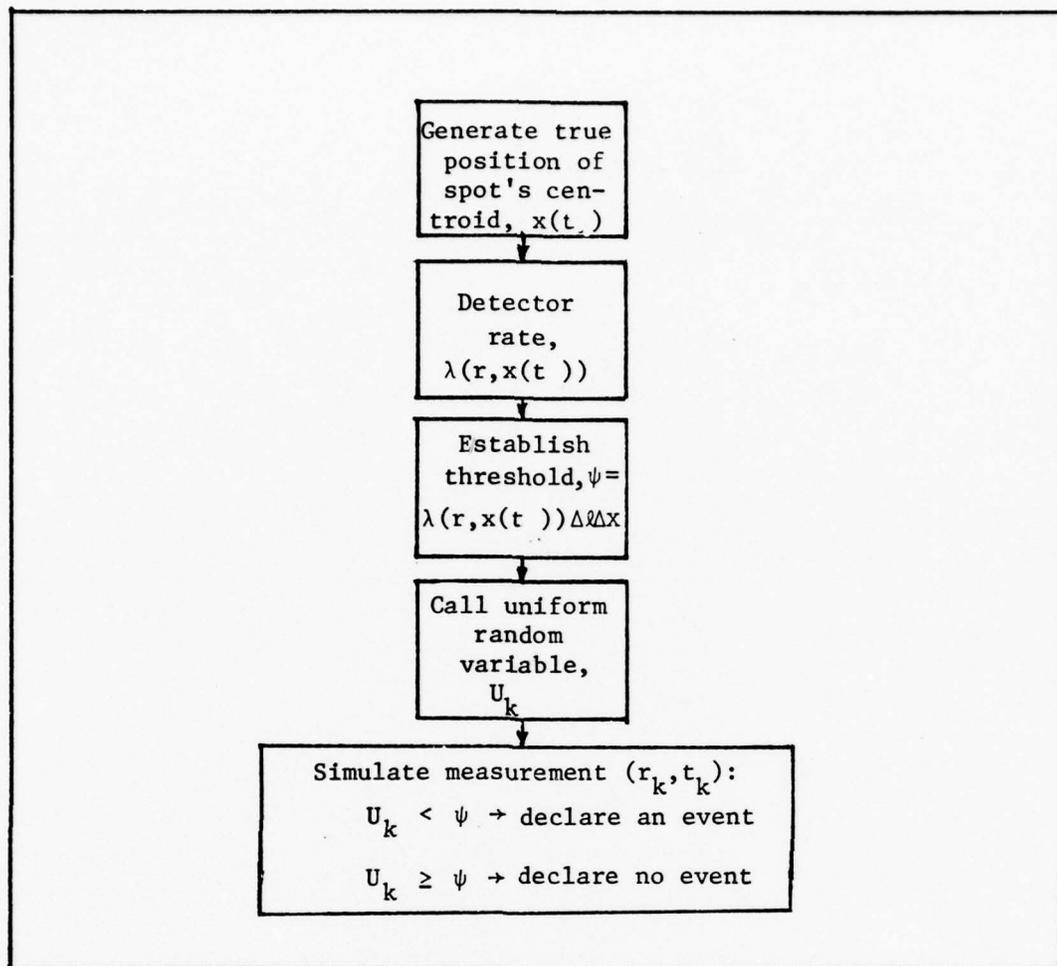
Figure 7. Flowchart of Binomial Method

one. Then, the probability of a value between zero and a number ψ, is simply ψ. So let $\psi = \lambda(\mathbf{r}, x(t)))\Delta\ell\Delta t$ establish a threshold to declare an event, where U_k is called from a uniform random generator (that is, $U_k < \lambda(r, \underset{\sim}{x}(t))\Delta t\Delta\ell$ means an event occurred at r on the detector surface).

The above module takes into account a different threshold and rate function when moving from one spatial point to another on the detector length. This method also includes effects of dark current (i.e., $\lambda(r, x(t)) = \lambda_g(r, x(t)) + \lambda_d$,

52

where $\lambda_g \triangleq$ Gaussian rate function $\lambda_d \triangleq$ dark rate function). The Binomial Method imposes a significant computational load in the simulation, as will be shown in a later section following the Conditional Method. Note that this method is an approximation for finite $\Delta \ell$ and Δt and that there is an issue of coarseness of the time and space discretization. This creates implementation and theoretical aspects which will be discussed in a later section.

Transformation Equation. The transformation equation is used to generate a desired random variable of a given probability distribution function by calling a uniform random generator. In this case, either an event location or event time is generated by calling a uniform random variable U which is distributed between zero and one with probability density height 1. The solution to this transformation problem is found in Ref. (12:184) and is given as

$$U = F_D(D) \tag{38}$$

or

$$D = F_D^{-1}(U) \tag{39}$$

where F_D is the desired distribution function, D is the desired random variable, F_D^{-1} is the inverse function of F_D. Fig. (8) graphically depicts this solution.

The random variable D represents either the random event location or time of the time-space point process. Similarly, F_D (or equivalently, the transformation function), corresponds

53

a) Desired probability distribution function, $F_D(D)$

b) Mapping to generate $F_D^{-1}(U)$
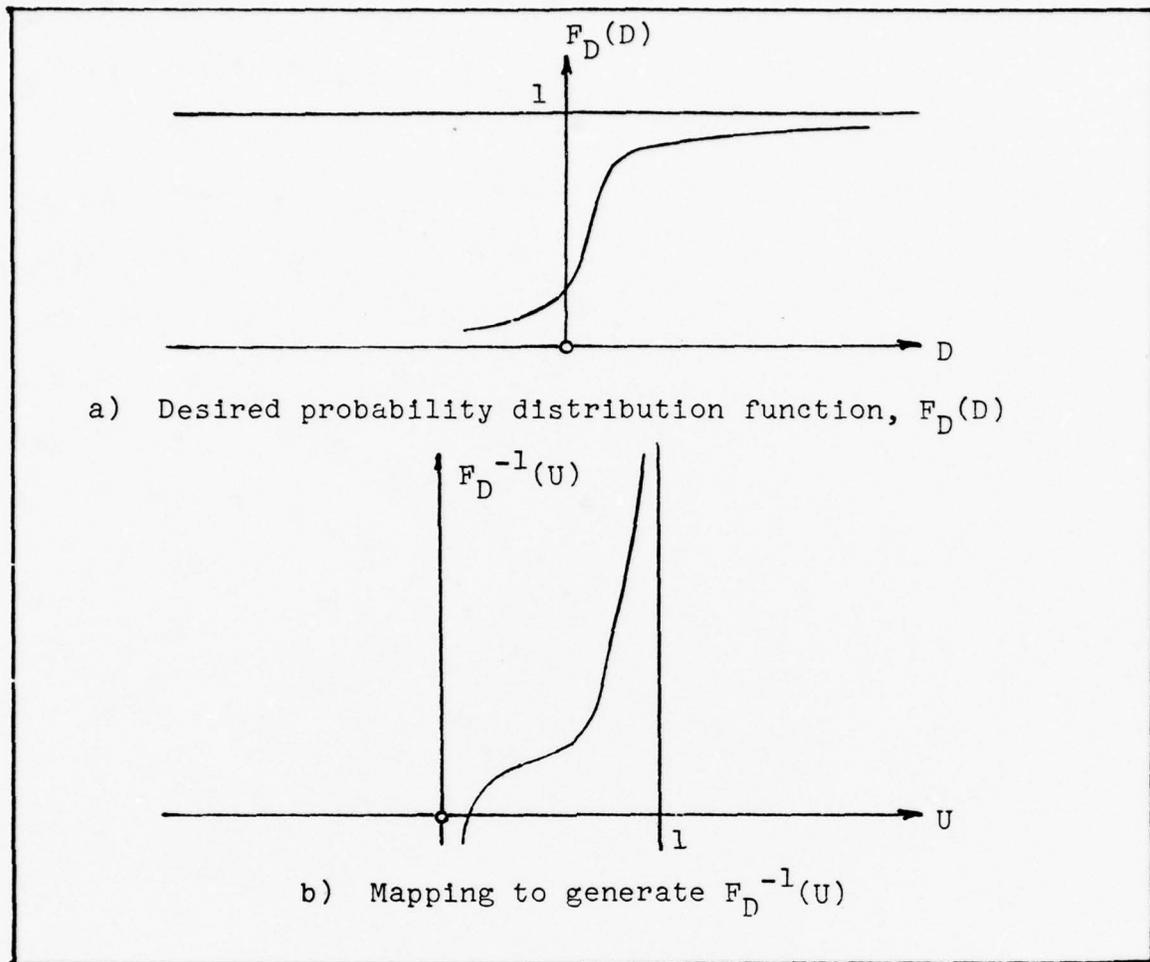
Figure 8. Generation of a Specified Probability
          Distribution Function (Copied from Ref. 12:184)

to either the event location or time probability distribution.
In this report, these probability distributions are derived from
the densities $f[r_k|t_k,n]$ and $f[s_k]$ where s_k is the interarrival
times of ordered events 1, (k=0, 1,2...,n), t_k is the event
times from a time orgin, and n is the number of events. These
densities are shown in the next subsection.

Conditional Method. To generate the time-space point
process events from the output process of the detector surface,

54

the number of events, the event times and event locations
are needed. Here, the probability density $f[\{r_k\},\{t_k\},n]$ is required
to characterize output. This joint density is employed to derive
the conditional densities of event times and event locations
conditioned on knowing certain variables, where the variables
are generated by actual implementation of the simulation. The
development of this section will involve some lengthy (but not
complex) derivations of expressions needed to generate $\{r_k\}$ and
$\{t_k\}$. Also, physical insights about the Snyder Filter's tracking
capabilities will be cited where appropriate.

In this study, the variables $I_0(t), \sigma(t), F(t), G(t)$,
$H(t)$, and $Q(t)$ are time-invariant. Then the average rate at
which photoelectrons appear on the photodetectors array (after
integrating the spatial dependence of the rate function,
defined in Eq. (8), over entire detector length) yields

$$\lambda_0 = \frac{\eta}{\eta f_0} I_0 \sqrt{2\pi} \ \sigma + \lambda_d \ell \tag{40}$$

where Eq. (14) was used, and ℓ is the detector length of array.
Eq. (40) has terms including the effects due to dark current;
however, in this relation, some assumptions were used, brought
forth earlier in the report. To explain the first term in Eq.
(40), it is assumed that the detector surface is large, such
that, with respect to the spot's Gaussian intensity profile,
the detector surface is infintely large. Thus, edge effects of
the spot's intensity profile on the detector may be ignored.
On the other hand, λ_d is finite which means that $\int_{-\infty}^{\infty} \lambda_d d\ell$

55

has an infinite value. Therefore, the integral, $\int_{-\frac{\ell}{2}}^{\frac{\ell}{2}} \lambda_d d\ell$
results in the second term of Eq. (40). The major point is
that the detector length must be large so that edge effects
between the spot and detector are neglible when the spot is
jittering about on the detector surface.

Consider the simulation of event times. It is known
that the interarrival time, denoted as s_k, has an exponential
density function (Ref. 3:60) given as

$$f[s_k] = \lambda_0 \exp(-\lambda_0 s) \tag{41}$$

To obtain $F_D[s_k]$, integrate Eq. (41) over the time continuum
between $-\infty$ and s_k. Using the transformation function from
Eq. (38), Eq. (41) yields

$$U_k' = F[s_k] = 1 - \exp(-\lambda_0 s_k) \tag{42}$$

where U_k' is a uniform random variable distributed between
zero and one. It then follows that $U_k = 1 - U_k'$ is also a
uniform variable with the same statistics. By straight forward
manipulation, Eq. (42) yields

$$s_k = -\frac{1}{\lambda_0} \ln U_k \tag{43}$$

To get the event times, just use a running sum. In this case,
with t_0 as the initial time the event times of the point process
with rate function λ_0 can then be simulated by the following
algorithm: call a uniform random variable, U_1; use Eq. (43) to
transform U_1 to s_1; assign s_1 as the time from t_0 to the first

photoconversion (i.e., $T_1 = t_0 + s_1$); call another uniform random variable, U_2; transform U_2 to s_2; assign s_2 as the time between the first and second photoconversion (i.e., $T_2 = T_1 + s_2$), and so forth (Ref. 3:62).

The next problem is finding an algorithm to generate the event locations, r_k's. This problem is solved by using densities $f[\{r_k\}, \{t_k\}, n]$ and $f[\{t_k\}, n]$ (see Appendix A for derivation). These densities describe the statistics of the output process from the photodetector surface which are expressed as

$$f[\{t_k\}, \{r_k\}, n]] = \exp(\iint_{Ad} {}_0^T \lambda(\alpha, \beta) d\alpha d\beta) \prod_{k=1}^{n} \lambda(t_k, r_k) \quad (44)$$

$$f[\{t_k\}, n] = \exp(-\int_0^T \lambda(\alpha) d\alpha) \prod_{k=1}^{n} \lambda(t_k) \quad (45)$$

where

$$\lambda(t_k, r_k) = \underbrace{\frac{n I_0}{f_0} \exp(\frac{-(r_k - \underset{\sim}{x}(t_k))^2}{2\sigma^2}) + \lambda_d}_{\lambda_g(r_k, t_k)} \quad (46)$$

$$\lambda(t_k) = \frac{n I_0}{f_0} \sqrt{2\pi} \ \sigma \ + \lambda_D \quad (47)$$

$$= \lambda_0$$

Note that $f[\{t_k\}, n]$ is the marginal density of $f[\{t_k\}, \{r_k\}, n]$ and that $\lambda(t_k)$ in Eq. (47) has no spatial or temporal dependence.

Also, the integrals in Eq. (44) and (45) will both yield $\lambda_0 T$. Thus Eqs. (44) and (45) respectively yield

$$f[\{t_k\},\{r_k\},n] = \exp(-\lambda_0 T) \prod_{k=1}^{n} \lambda(t_k,r_k) \qquad (48)$$

$$f[\{t_k\},n] = \exp(-\lambda_0 T) \prod_{k=1}^{n} \lambda(t_k) \qquad (49)$$

With the help of Bayes' rule, the conditional probability density function of event locations conditioned on knowing the event times and the number of events is given as

$$f[\{r_k\}|\{t_k\},n] = \frac{f[\{t_k\},\{r_k\},n]}{f[\{t_k\},n]}$$

$$= \prod_{k=1}^{n} \frac{\lambda_g(t_k,r_k) + \lambda_d}{\lambda_0}$$

or

$$f[\{r_k\}|\{t_k\},n] = \left[\frac{\lambda_g(t_1,r_1) + \lambda_d}{\lambda_0}\right]\left[\frac{\lambda_g(t_2,r_2) + \lambda_d}{\lambda_0}\right]$$

$$\cdot \left[\prod_{i=3}^{n} \frac{\lambda_g(t_k,r_k) + \lambda_d}{\lambda_0}\right]$$

$$= f_1[r_1|t_1,n]\, f_2[r_2|t_2,n]$$

$$\cdot \prod_{k=3}^{n} f_k[r_k|t_k,n] \qquad (50)$$

58

where Eqs. (48) and (49) were used. The above relation says that the joint conditional density $f[\{r_k\}|\{t_k\},n]$, can be derived by forming the product of each of the densities at the time-space coordinates, (t_k, r_k). Thus, for n=1 at (r_k, t_k), the conditional density of r_k given t_k is expressed as

$$f[r_k|t_k] = \frac{\lambda_g(t_k, r_k)}{\lambda_0} + \frac{\lambda_d}{\lambda_0} \tag{51}$$

Before proceeding further, it is useful to check if Eq. (51) is normalized (i.e., $\int_{-\infty}^{\infty} f[\alpha_k|t_k]d\alpha_k=1$). That is,

$$\frac{1}{\lambda_0} \int_{-\infty}^{\infty} \frac{\eta}{f_0} I_0 \exp\left(\frac{-(\alpha_k-x)}{2\sigma^2}\right)d\alpha_k + \frac{1}{\lambda_0} \int_{-\frac{\ell}{2}}^{\frac{\ell}{2}} \lambda_d d\alpha_k$$

$$= \frac{1}{\lambda_0} \left(\frac{\eta}{hf_0} I_0 \sqrt{2\pi} \sigma + \lambda_d \ell\right) = 1$$

Again, it is assumed that the detector is large such that the limits in the integral of the first term on the left side of Eq.(52) are approximately valid.

To get the probability distribution function $F[r_k|t_k]$ integrate $f[r_k|t_k]$ over the detector plane from $-\infty$ to r_k for the first term and from $-\frac{\ell}{2}$ to r_k for the second term in Eq. (51). The result is expressed as

$$F[r_k|t_k] = \frac{1}{\lambda_0}[\text{erf } v + \lambda_d(r_k+\tfrac{\ell}{2})] \tag{53}$$

where

$$\text{erf } v \triangleq \frac{2}{\sqrt{\pi}} \int_0^v \exp(-Z^2) dZ$$

$$v \triangleq \frac{r_k - x(t_k)}{\sqrt{2}\ \sigma}$$

Using the transformation equation defined in Eq. (39) and after algebraic manipulation, Eq. (53) yields

$$\text{erf } v + (\frac{2}{\sqrt{\pi}} A_1)\ v = 2A_2\ U_k - 1 + 2A_3\ x(t_k) + \frac{\ell}{2} \qquad (54)$$

where

$$A_1 \triangleq \frac{\lambda_d}{\dfrac{\eta I_0}{hf_0}}$$

$$A_2 \triangleq \frac{\dfrac{\eta I_0}{(\ f_0\ \sqrt{2\pi}\ \sigma + \lambda\ \ell)}}{\dfrac{\eta I_0}{hf_0}\ \sqrt{2\pi}\ \sigma}$$

$$A_3 \triangleq \frac{\lambda_d}{\dfrac{\eta I_0}{hf_0}\ \sqrt{2\pi}\ \sigma}$$

Eq. (54) is used to generate the event locations at the event times given the variables defined in this expression. The event locations are generated by a table lookup method with linear interpolation and is performed as follows:

    (1) Evaluate and obtain all necessary variables, i.e., A_1, A_2, A_3, $\frac{\ell}{2}$, U_k, and $x(t_i)$,

(2) For a specific range of values for v, solve the expression on the left-hand side of Eq. (54), and store the results in a table (i.e., computer memory),

(3) Derive a random variable s by calling a uniform random variable, U_k, distributed between (0,1) and by obtaining the true position of the centroid, $x(t)$, from the truth model simulation. Given these values, the right-hand side of Eq. (54) is evaluated and the random variable realization s is obtained,

(4) To find v, obtain s by a table lookup with linear interpolation. In the table the increment of v is equal to 0.025, where erf v is piecewise linear for this approximation; thus, linear interpolation is used

(5) Upon finding v, it then follows that the event location is evaluated by the following expression:

$$r_k = (\sqrt{2}\ \sigma)v + x(t_k)$$

This paragraph completes the algorithm of generating the event locations.

Since Eq. (54) governs the generation of event locations, it is useful to examine this relation for limiting cases (i.e., $\lambda_d = 0$ and $\lambda_g(t_k, r_k) = 0$). This investigation will provide insights into the tracking capabilities of the Snyder Filter.

$\lambda_d = 0$ implies $A_1 = A_3 = 0$ and $A_2 = 1$. Eq. (54) is then manipulated into a convenient form as

61

$$U_k = 1/2\left(1 + \operatorname{erf}\left(\frac{r_k - x(t_k)}{\sqrt{2}\ \sigma}\right)\right) \tag{55}$$

$$U_k = F_D(r_k | t_k) \tag{56}$$

where Eq. (38) is implemented in the last expression. Corresponding to $F_D(r_k | t_k)$ is $f(r_k | t_k)$ which is given as

$$f[r_k | t_k] = \frac{1}{\sqrt{2\pi}\ \sigma} \exp\left\{\frac{-1}{2\sigma^2}[r_k - x(t_k)]^2\right\}^2 \tag{57}$$

The event locations are generated by calling a Gaussian random variable of mean $x(t_k)$ and variance σ, at the event times. Note that the intensity profile of the spot being Gaussian-shaped leads to a result that the event locations are generated as Gaussian random variables conditioned on knowing the event times. The Snyder Filter is based on the assumption of a Gaussian-shaped profile of the spot. Thus, for $\lambda_d \neq 0$, A_1 "smears" out the Gaussian information as seen in Eq. (54).

For $\lambda_g(r,t) >> \lambda_d$, the effect of dark current is examined by doing a perturbation analysis. This analysis is performed by factoring out the larger of the two rate functions in Eq. (47). Then the natural logarithm of the left and right hand side of Eq. (47) becomes

$$\ln f(\{r_k\} | \{t_k\}, n) \approx \sum_{k=1}^{n} \ln\frac{\lambda_d}{\lambda_0} + \sum_{k=1}^{n} \frac{\lambda_g(r_k, t_k)}{\lambda_0}$$

where the approximation $\ln\left(1 + \frac{\lambda_g(r_k, t_k)}{\lambda_d}\right) \approx \frac{\lambda_g(r_k, t_k)}{\lambda_d}$ and

62

$\dfrac{\lambda_g(r_k,t_k)}{\lambda_d} << 1$ is used. The exponential of the left and right hand side of the above equation yields

$$f[\{r_k\}|\{t_k\},n] \cong \prod_{i=1}^{n} \frac{\lambda d}{\lambda_o} \exp\left(\frac{\lambda_g(r_k,t_k)}{\lambda_o}\right) \qquad (58)$$

Again, Eq (58), can be derived by forming the product of the above individual densities at (r_k,t_k). Then, setting $\lambda_g(r_k,t_k) = 0$ in Equation (58) and applying this condition to Eq (40), Eq (58) yields

$$f[r_k|t_k] \cong \frac{1}{\ell} \quad ; \quad -\frac{\ell}{2} < r_k < \frac{\ell}{2} \qquad (59)$$

where n = 1. Intuitively, Eq (58) is reasonable, since at any given time a dark event can occur in any differential length on the detector surface with equal probability.

In summary, for signal-limited conditions, the event locations are called as a Gaussian random variable with mean $\underline{x}(t_k)$ and variance σ^2; for dark-limited conditions the event locations are generated as uniform random variables distributed in $(-\frac{\ell}{2},\frac{\ell}{2})$ with probability density height $\frac{1}{\ell}$; and for generating event locations other than the limiting cases, Eq (54) is implemented. Finally, the generation of event locations is called at the event times given by Eq (43).

Aspects of Implementing each method. The Binomial Method has two restrictions for its algorithm to work properly: one, is $\lambda[r,x(t)]\Delta\ell\Delta t\leq 1$ and the other, is $\Delta\ell\Delta t<<1$ where for implementation purposes $\Delta t \triangleq$ sample time or observation time

63

interval. Because of these conditions, several practical and theoretical problems arise. One is a resolution problem. The Snyder Filter assumes that its position-sensitive measurements are from an infinitesimally fine partitioned detector array. No matter how small "$\Delta\ell$" is, there is still a problem in extracting the "exact" location of the event. For this report, the detector array is divided into equal increments of $\Delta\ell$, and the value of the rate function at the center of each incremental length, is chosen to evaluate the probability, $\lambda(r,x(t))\Delta\ell\Delta t$. This procedure determines where an event (if any) occurred on the detector surface. This method is depicted in Fig (9). Note that Δt is the observation increment during the simulation time. From a practical point of view, this method consumes a large amount of computer time; at each time increment of observation, $\Delta\ell$ is scanned across the detector surface from one end to the other.

In the Binomial Method, a theoretical point that should be considered is the question of "simultaneous" events occurring on the surface. The Snyder Filter assumes that one event is observed at any one time. With this simulation method, there is a probability that more than one event can occur. Even if $\Delta t << 1$ and $\Delta\ell << 1$, there is still a probability (though small) that simultaneous events can occur at different locations. It was mentioned previously that scanning the detector surface for small $\Delta\ell$ results in a large comsumption of computer time at each time increment. Thus, if a smaller Δt or $\Delta\ell$ is chosen to reduce the probability of simultaneous events, this will result in a larger consumption of computer resources.

64
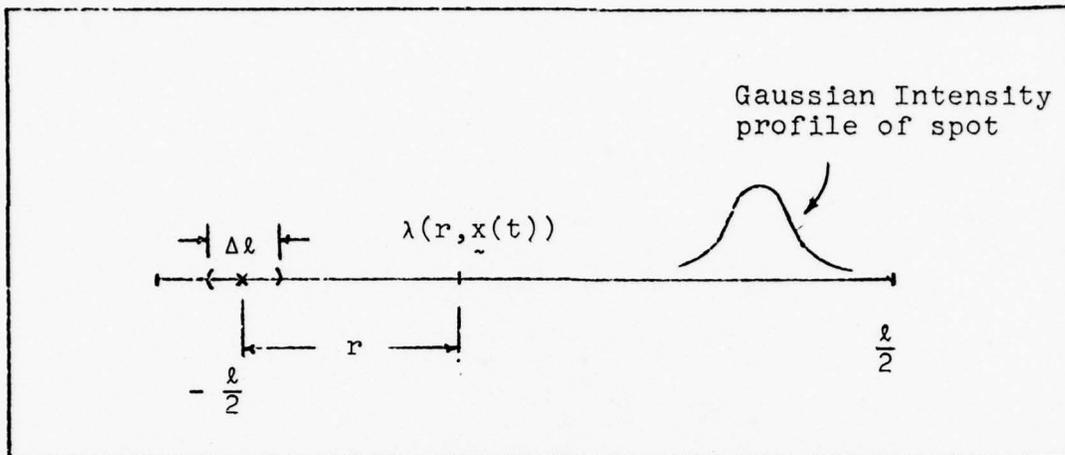
Fig. 9.   Resolution problem in Binomial Method

Furthermore, when dark current is included in the simulation, the probability of simultaneous events grows substantially. To reduce the amount of computer resources (and to observe the errors generated by the Snyder Filter between event times), the rate functions are chosen to have a low value. This enables the use of a larger sample period, Δt, which also implies a higher probability of simultaneous events. It is evident now that the Binomial method has implementation problems. In spite of these problems this method is performed in the computer, and the results are shown in the next chapter.

In terms of implementation, the Conditional Method is a better algorithm, since the event locations are called at the event times; (i.e., the problem of simultaneous events is precluded). In the simulation it is desirable to sample the errors committed by the Snyder Filter between successive event times. In this case, it is convenient to break up the simulation time interval (observation time) into time increments, Δt. Δt will denote the sample period in order to sample the

65

error process (as done in the Binomial Method). When the event times are generated, there is a situation that the interarrival times, w_k's, can occur smaller than the sample period (i.e. $s_k < \Delta t$). This situation is shown in figure (10).

Performance Analysis

The objective of the performance analysis is to characterize the error process statistically. This is done by generating many samples of the error process through Monte-Carlo simulation. The simulation uses random number generators and shaping filters to generate the random errors. One objective in the simulation is to portray filter performance capabilities when parameters that describe the problem are varied. First, consider the test cases without dark current. Here, the signal rate function or spot's jitter are changed and, thus, filter capabilities can be evaluated. Note that changing the signal rate varies the average number of photons detected per coherence time; similarly, changing the spot's jitter causes variations in the Y parameter. In this study, the test cases do not meet sufficient conditions which are $X >> 1$ and $Y << 1$; therefore, the results should provide an indication of which bound is close to actual performance. A plot which contains the sample variance and average filter variance is a useful output to make this comparison.

Second, the measurement model (of which the Snyder Filter is based) is altered by presenting dark events in the measurement process. The above plot is again a useful output to evaluate the filter's sensitivity to dark measurements.

In this study, a final objective of the performance analysis is to change the filter structure, through ad hoc methods, when

66
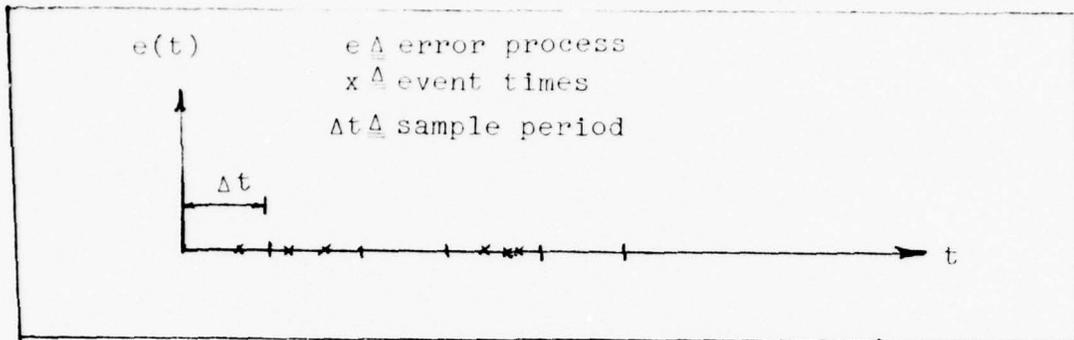
Fig 10. Sampling the Error Process with Fixed Sample Period
(Conditional Method)

satisfactory performance cannot be obtained in the presence of
dark events. Again, the above plot containing the variances
would be useful in providing insights of how to alter the
filter structure.

As previously mentioned, the error process is characterized by
generating many sample runs of this process. The appropriate
number should be chosen sufficiently large, such that additional
sample runs does not appreciably change the sample statistics.
The last statement is an illustration of the weak law of large
numbers: let $E[\bar{e}] = \bar{e}$ denote the ensemble average of the
error process $\bar{e}$. Also, let $<\bar{e}>_N$ denote the empirical average for
N runs. When N is large, the weak law states with high
probability that $<\bar{e}>_N$ is close to the number $\bar{e}$. In mathematical terms,

$$P[|<\bar{e}>_N - \bar{e}| > \epsilon] = \frac{\sigma_{\bar{e}}^2}{N\epsilon^2}$$

where $<\bar{e}>_N \triangleq \frac{1}{N} \sum_{j=1}^{N} e_j$ sample mean error, $N \triangleq$ number of
statistical independent random variables, e_j with mean, $\bar{e}$
and variance $\sigma_{\bar{e}}^2$, and ϵ is an arbitrary constant. The

67

weak law is a measure of the confidence of inferring information about the actual statistics of the error process (Ref 11:96). Thus, as the number of sample runs increases, the confidence in characterizing the error process accurately increases.

Sample statistics. The statistics computed at each time point are the calculated mean error $\bar{e}$, calculated variance σ_e^2, and the ensemble average of the filter variance terms $\overline{PF}$. These calculations are performed over the ensemble of runs N, for each time point t_m. These expressions are

$$\langle e(t_m) \rangle_N = \frac{1}{N} \sum_{j=1}^{N} [\hat{x}(t_m) - x(t_m)] \tag{60}$$

$$\sigma_e^2(t_m) = \frac{\sum_{j=1}^{N} [x \; t_m) - x(t_m)]^2}{N-1}$$

$$- \sum_{j=1}^{N} \frac{N[\langle e(t_m) \rangle_N]^2}{N-1} \tag{61}$$

$$\overline{PF}(t_m) = \frac{1}{N} \sum_{j=1}^{N} PF(t_m) \tag{62}$$

The above relations are found in Ref (13). Note that the time point t_m are not event times but are the sample times for a fixed sample period Δt. From Eq (61), the variance $\sigma_e^2(t_m)$ is divided by (N-1) instead of N, which results in an unbiased estimator of variance (Ref 13).

Plots. Mentioned in previous paragraphs is that a useful output in the analysis is a plot of the average filter estimate of the error variance, along with the corresponding computed

68

sample (or true) variance for all times of interest. This output
is used to compare if the sample variance of the error process
matches with the filter's average error variance. In other
words, this output should indicate how well the filter is
estimating its errors with the actual errors it commits; and
to compare with the bounds on MSE for test cases where dark
current is set equal to zero.

A second useful output on this analysis is a plot of the
mean error versus time. Along with this plot are the standard
deviation bounds plotted above and below the mean error. This
plot is useful if the Snyder Filter provides the desired accuracy.
The two plots that were just described, are used to analyze the
average error performance of the Snyder Filter. For all
cases, these plots are used to determine whether the actual
errors committed by the filter are sufficiently small to yield
good tracking performance.

Now consider plots of individual sample of the error
process. For no dark current, this plot is used to observe how
the Snyder Filter minimizes its errors. When dark current is
represented in the simulation, then the Snyder Filter sensitivity
to the dark measurements can be observed. Along the zero axis
of this plot are the event times to indicate clearly how the
filter updates its estimate at the event times. Also, a plot
of the sample filter variance is also of interest in order to
observe how the Snyder Filter propagates the uncertainty of its
estimate in time.

Test cases. During this study, different cases using different values of X and Y performance parameters are considered. The test cases are given in tables (1) and (2). All test cases are performed where there is a relative separation between the bounds. Here, the bound lying closer to actual performance can be determined. For the test cases in Table (1) and (2), the dark current is set equal to zero.

When dark current is presented in the simulation separate test cases are performed. The measurement process of the Snyder Filter is altered by this addition of dark current. For these test cases, a sensitivity parameter is defined to be the average signal-to-noise counts ratio which is expressed as

$$\frac{S}{N} = \frac{\frac{nI_o}{hf_o} \sqrt{2\pi} \; \sigma}{\lambda_d \cdot \ell}$$

Here, the degree of degradation in filter performance is observed as the average signal-to-noise counts ratio is varied. Because of the time constraints imposed in the study, limited cases for $\lambda_d \neq 0$ are examined.

Summary

The chapter developed the necessary algorithms needed to perform this study. These algorithms consists of simulating the spot's dynamics, and the generation of position-sensitive measurements (i.e. time-space point process). The chapter concluded with a general discussion on performance analysis, where it consists of sample statistics, plots, and test cases needed to investigate the tracking capabilities of the Snyder Filter.

70

Performance Parameter	
X	Y
5.0	0.1
5.0	0.25
5.0	1.0

Table 1.   Test Cases for Binomial Method
(without dark current)

Performance Parameter	
X	Y
5.0	0.5
5.0	1.0
5.0	5.0
5.0	10.0
1.0	0.5
1.0	1.0
1.0	5.0
1.0	10.0
0.5	5.0
10.0	5.0
15.0	5.0
20.0	5.0
0.5	1.0
10.0	1.0
15.0	1.0
20.0	1.0

Table 2.   Test Cases for Conditional Method
(without dark current)

IV Results

This chapter contains results from the Monte Carlo simulation.
The simulation will indicate how well the actual performance can
be described, when compared to the upper and lower bounds. Thus,
various values on the set of parameters, derived from the bounds,
are used to determine which bound is closer to actual performance.
The bounds are valid only when the measurement is perfect (i.e.,
there are no dark or background noise events). When dark current
is included in the simulation, the results should show the degree
of degradation in filter performance. Finally, to compensate for
the effects of dark current and to improve tracking performance,
different methods in changing the filter structure are exploited.

Mean and Variance Convergence

To determine the appropriate number of sample runs, two
plots are generated. One plot shows the sample mean error versus
the number of runs, and the other plot shows the sample variance
of error process versus the number of runs. These plots are shown
in Fig. (11) for the Binomial Method. On each of the two plots
are four curves corresponding to four arbitrary time points
(as indicated in the figure) during the simulation time.
By the weak law of large numbers, the mean and variance of the
error process should converge to a steady value such that the
variance or mean of the error process does not change much in
magnitude with each addition of sample runs. Based on this
reasoning and the results from the plots, fifty sample runs are
selected. This number of runs will also be used in the conditional
method. However, it does not appear that fifty runs are sufficient

72
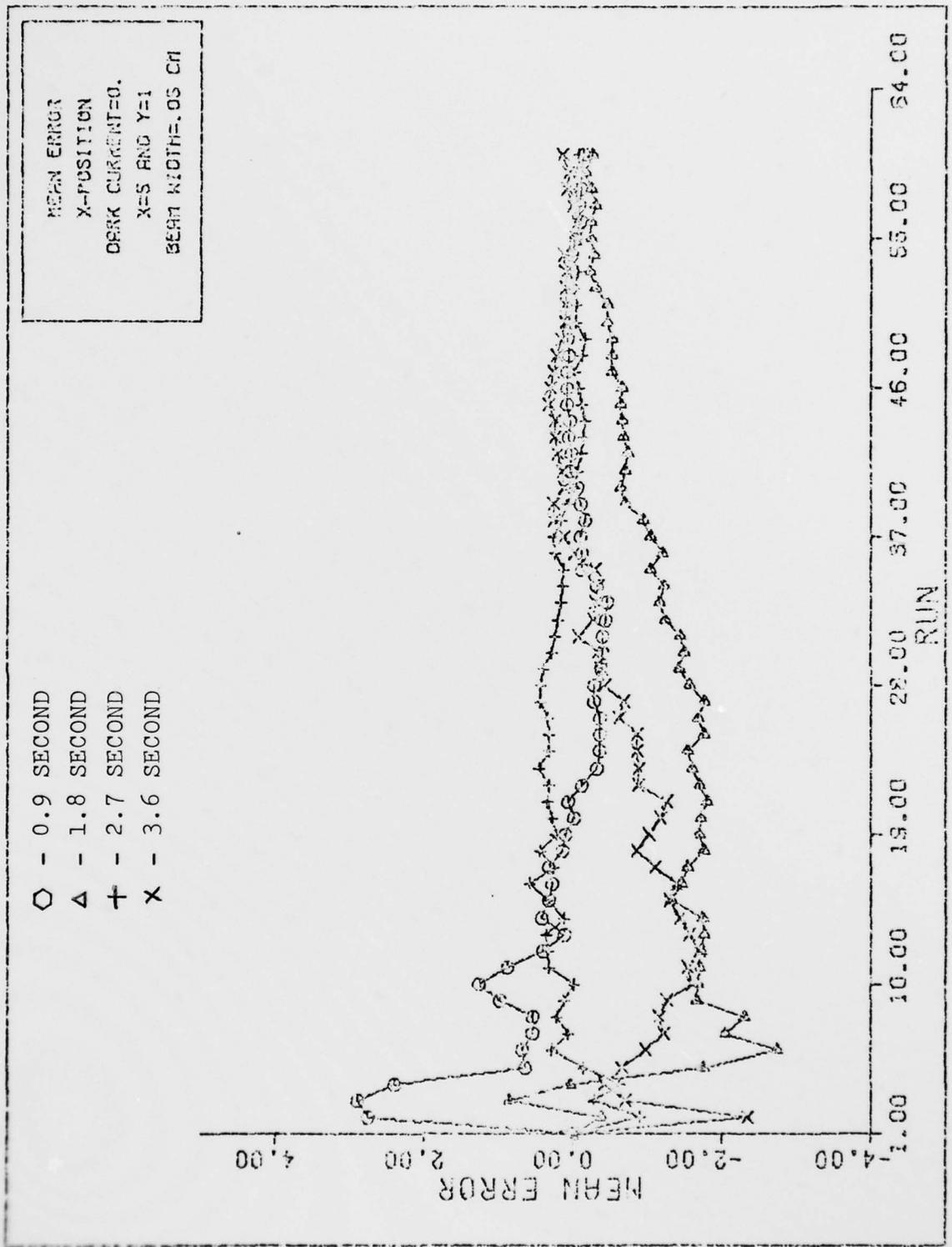
Figure 11a MEAN CONVERGENCE

73
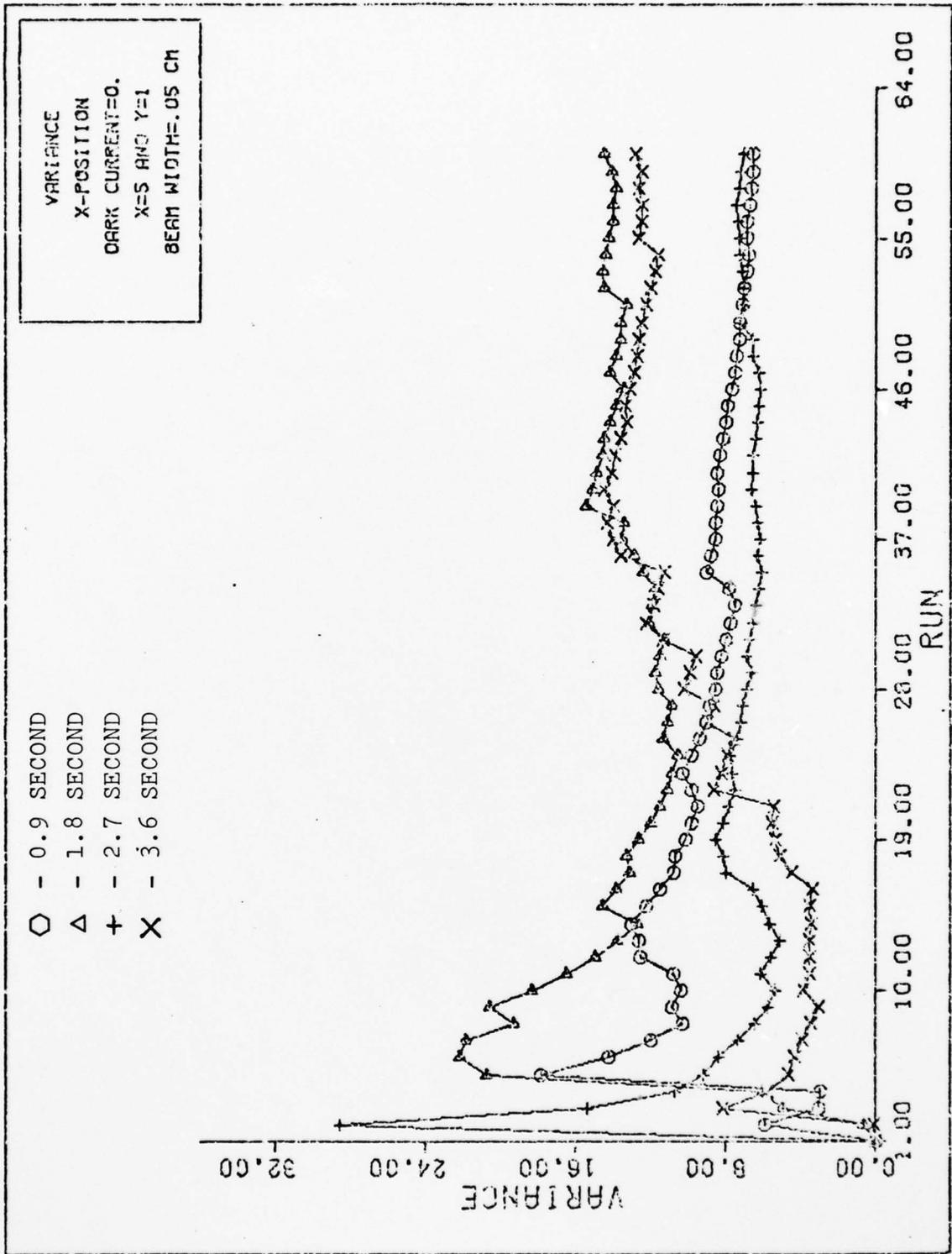
Figure 11b VARIANCE CONVERGENCE

74

to ensure convergence for all the test cases listed in Tables
1 and 2. That is, the performance parameters, X and Y, which are
defined in Chapter II as

$$X = \left(\frac{nI_0\sqrt{2\pi}}{hf_o} \right) (\tfrac{1}{2}\tau_c) \underline{\triangleq} \text{ average number of electrons}$$
occurring in one-half the
coherence time of spot's
dynamics $(\lambda_d=0)$ (63)

and

$$Y = \frac{\sigma_x^2}{\sigma^2} \underline{\triangleq} \quad \text{spot's jitter normalized by the square of}$$
the beam width (64)

have different values. The plots shown in Fig. (12) provide
an additional illustration of confidence in the sample variances.
They are plots of the variances of the error process (calculated
from sample statistics) for different numbers of sample runs.
On the true variance curves are symbols, denoted as "x", to
distinguish *from the filter* variance curve. Along with the curves
are the upper and lower bound on MSE for the given set of parameters
X and Y. These plots show that as the number of simulation runs
increases, the fluctuations of the true and filter variances
decrease. The plots also indicate that the fluctuations should
settle to a value that is between the upper and lower bounds (as
expected). Hence, the number of runs indicate the amount of
confidence that the sample and average filter sample variances
approach the corresponding true variances. Although fifty
simulation runs may not be appropriate for all the test cases,
it does provide an indication of filter performance.

75
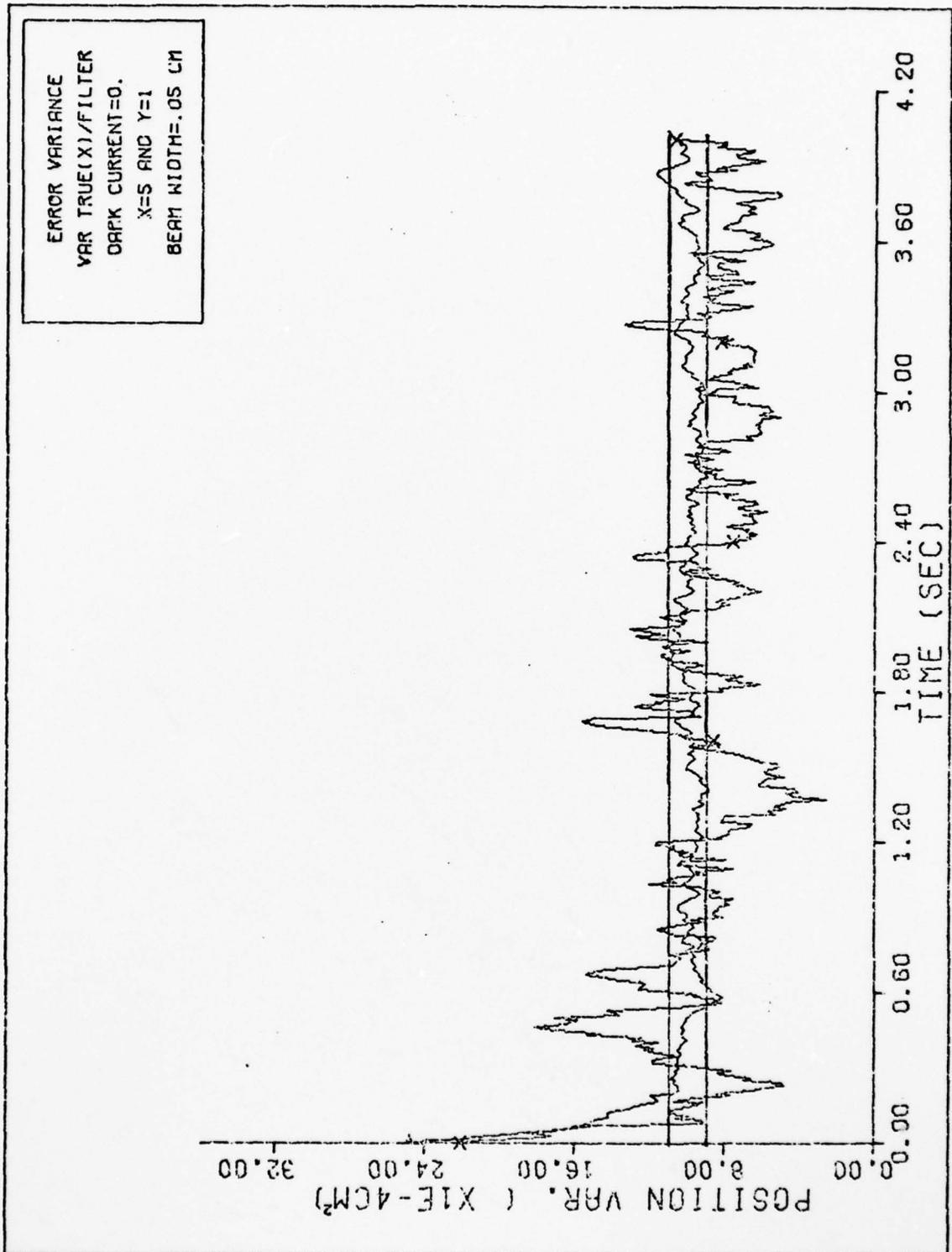
Figure 12a X-POSITION VARIANCES

(20 SAMPLE RUNS)
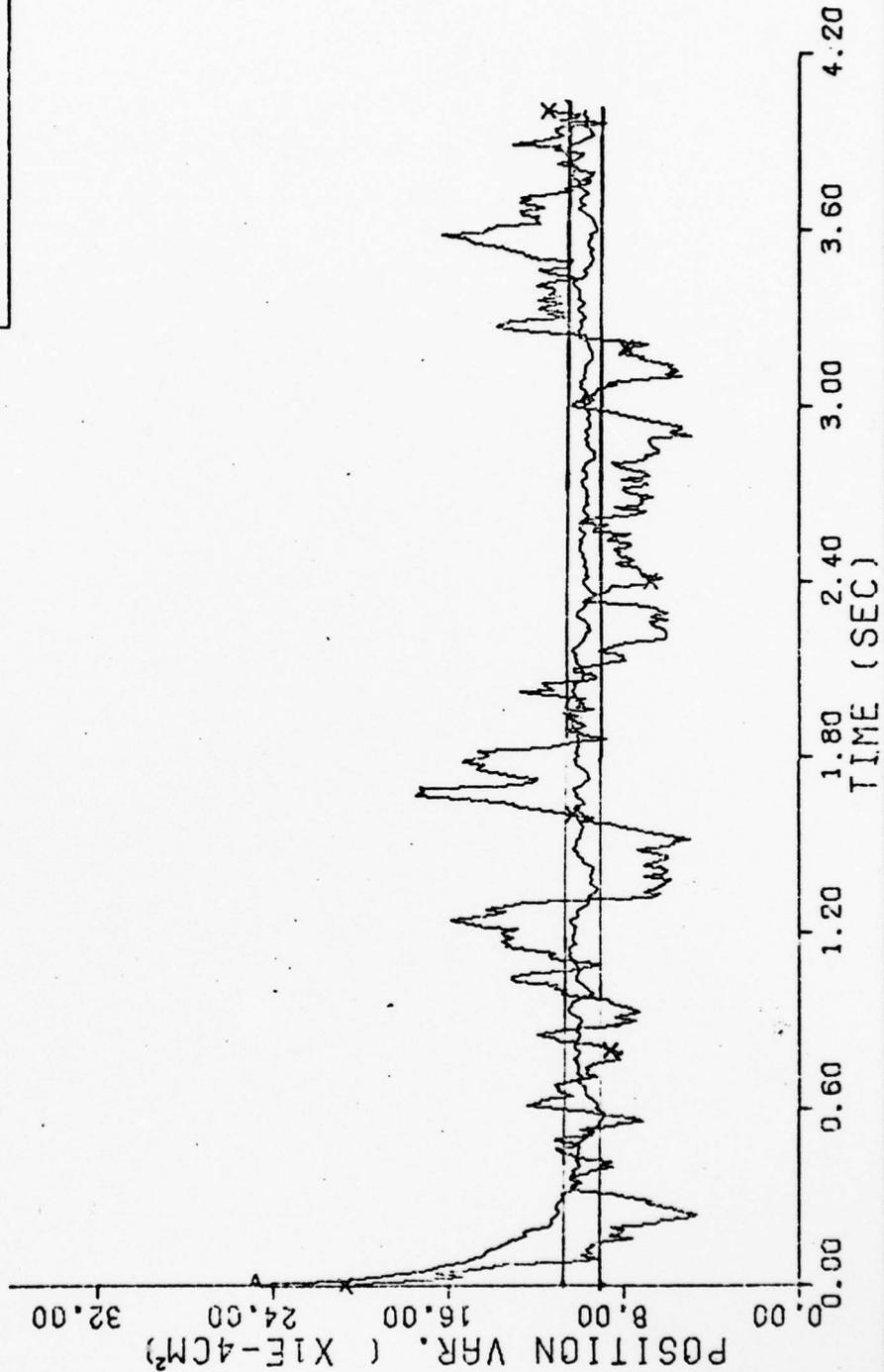
Figure 12b X-POSITION VARIANCES
(40 SAMPLE RUNS)

77
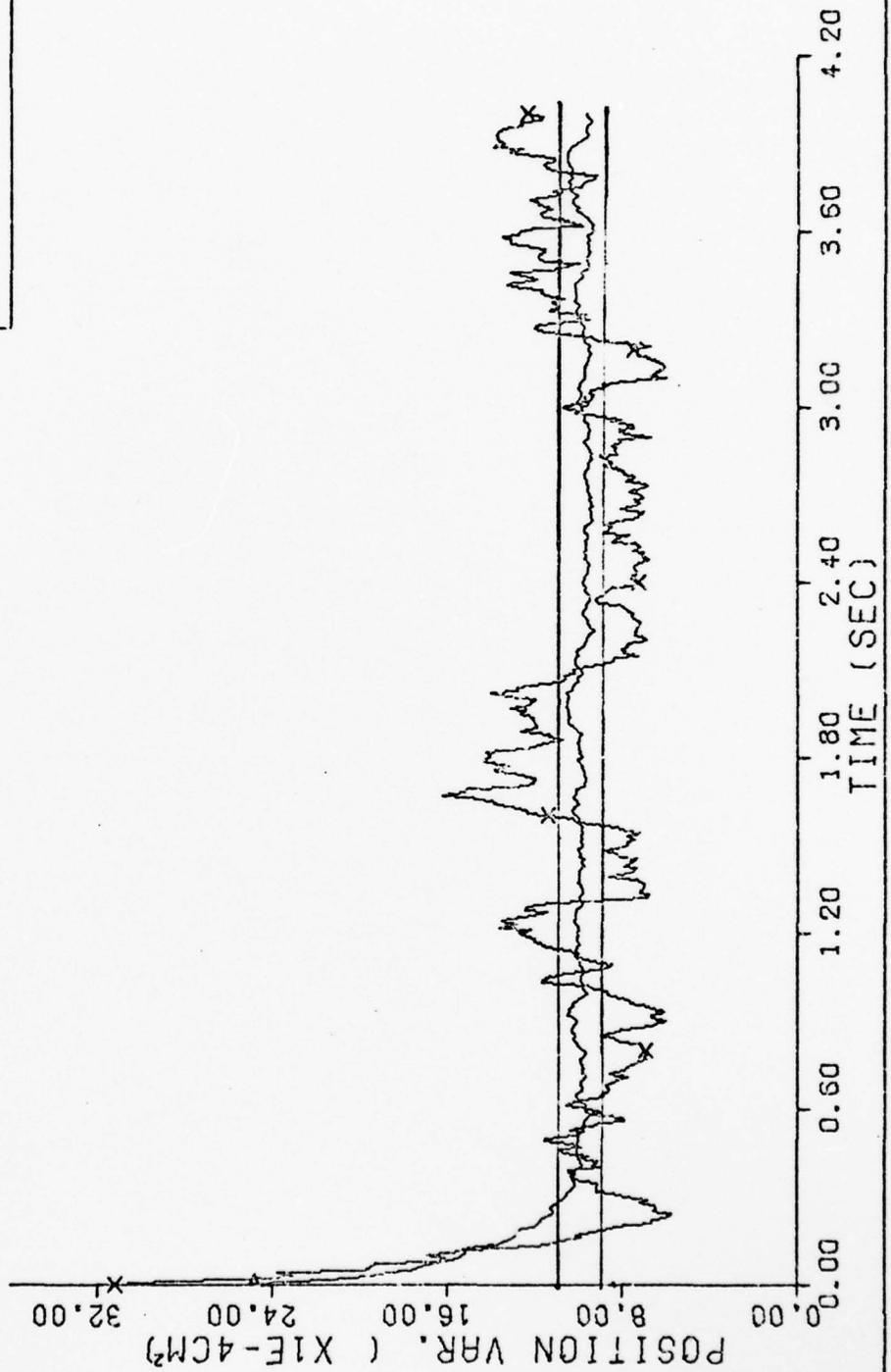
Figure 12c X-POSITION VARIANCES

(60 SAMPLE RUNS)

78
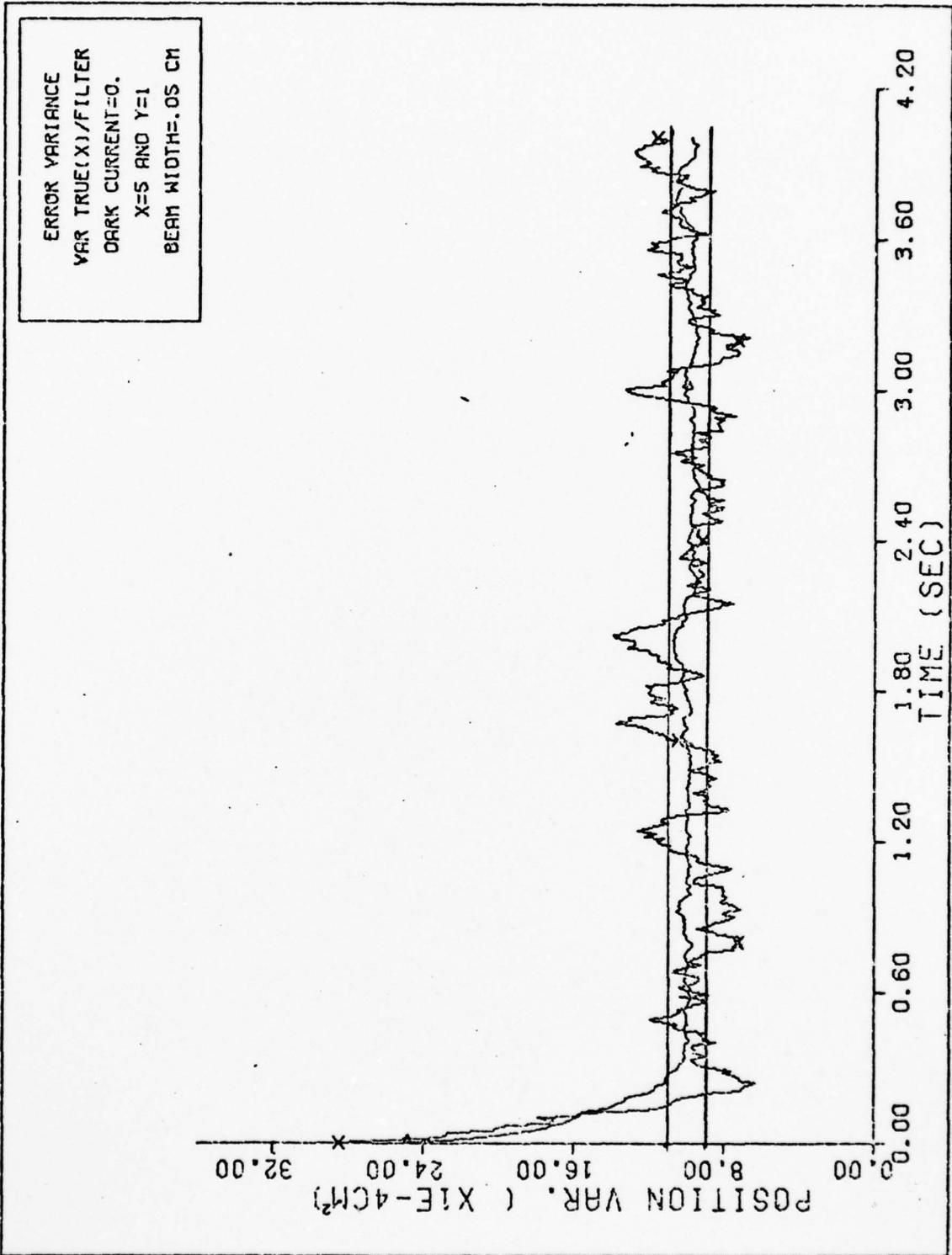
Figure 12d X-POSITION VARIANCES
(80 SAMPLE RUNS)

79
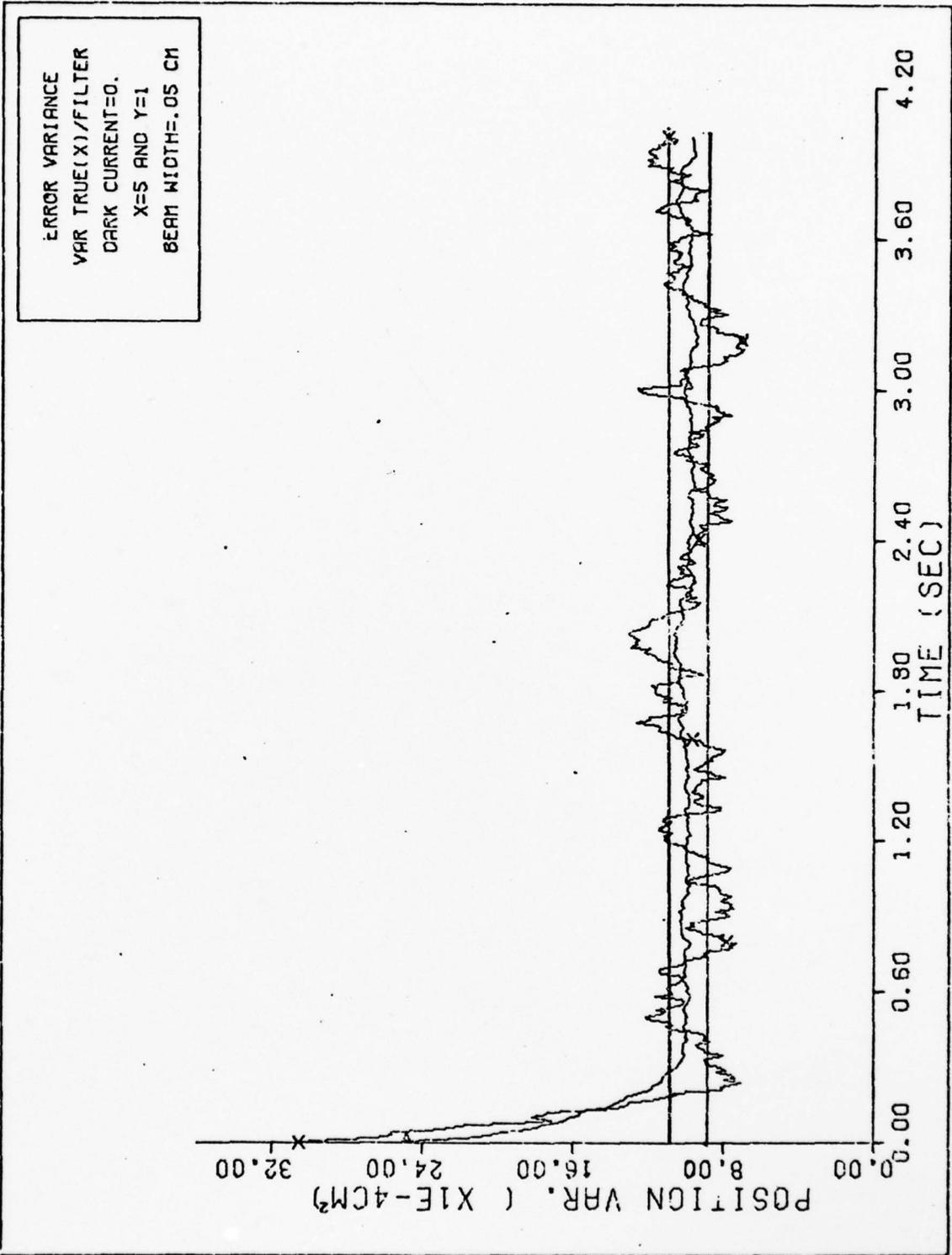
Figure 12e X-POSITION VARIANCES

(100 SAMPLE RUNS)

80

Binomial Method (without dark current)

For the Binomial Method, it was mentioned in the previous chapter that simultaneous events can be generated at the sample period, Δt. This situation would occur often if the observation area, $\Delta t \Delta \ell$, is not sufficiently small. In addition, a low rate function, $\lambda(t,r)$, is selected to observe the propagation of the conditional mean and variance of the Snyder Filter and to reduce the amount of computer resources by this method. During this part of the study, it is found that simultaneous events did occur even for a rate function of 10 events per second with a sample period of 0.01 second. For X=5, simultaneous events occurred on the average of two to three times per sample run. To avoid this problem, a software routine is developed to randomly pick one of two simultaneous events. Note that this routine ruins the statistical properties of the point process.

Figs. (13a, 14a) and (13b, 14b) are plots of the sample error and sample filter variance, respectively, for the parameters X=5 and Y=1. Note on the time axis of each plot are the event times. In Figs. (13a) and (14a), the errors decrease as events appear frequently in time (as expected). Similarly, in Figs. (13b) and (14b), the conditional variance is updated at the event times, otherwise the variance grows in time when an event is not observed. Note, the filter variance is "driven" to a smaller value as more events appear frequently in time. If the filter does not observe many measurements, the filter will weigh the next measurement heavily, driving the filter variance to a lower value. If many previous measurements appear frequently in time, this implies

81
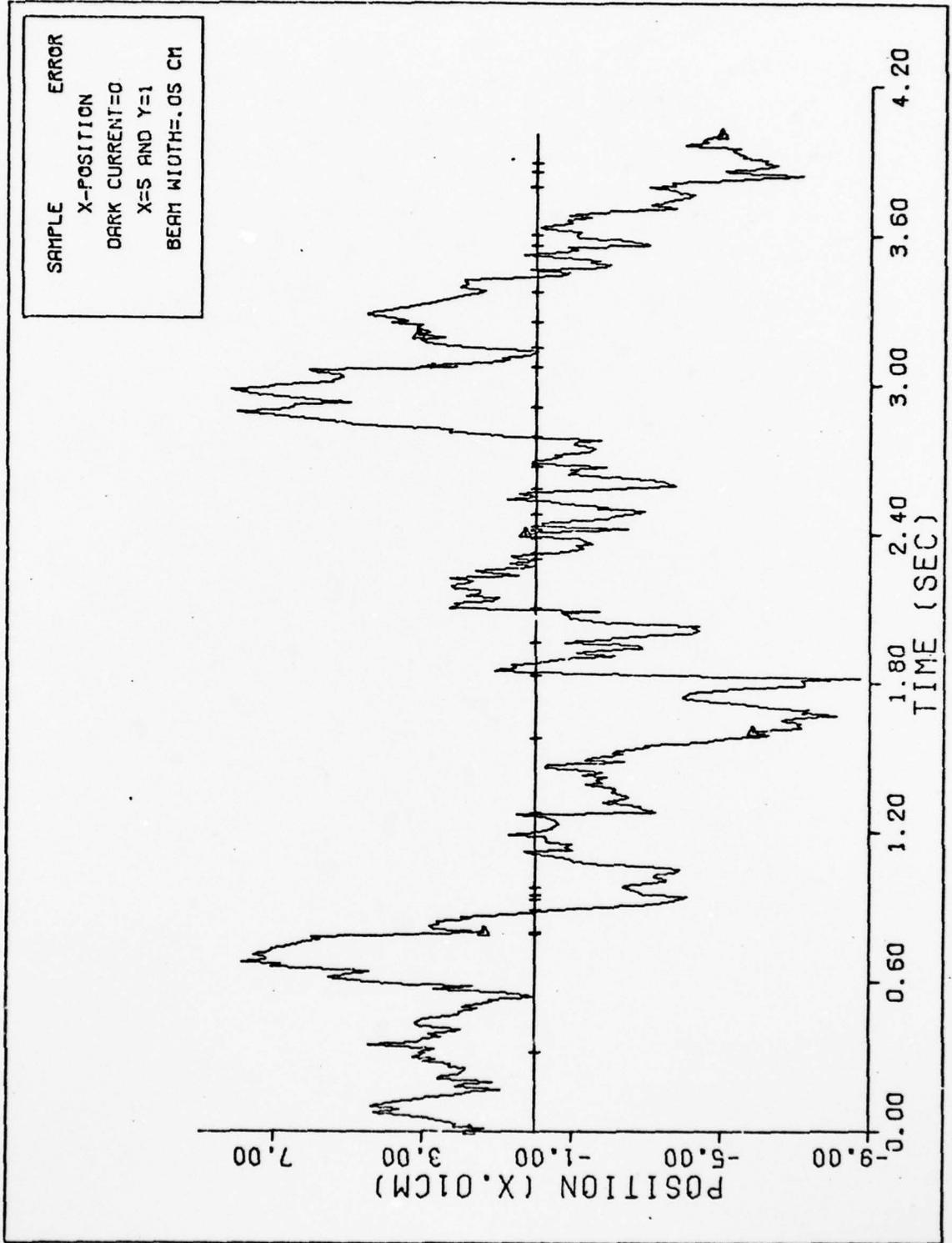
Figure 13a X-POSITION

82
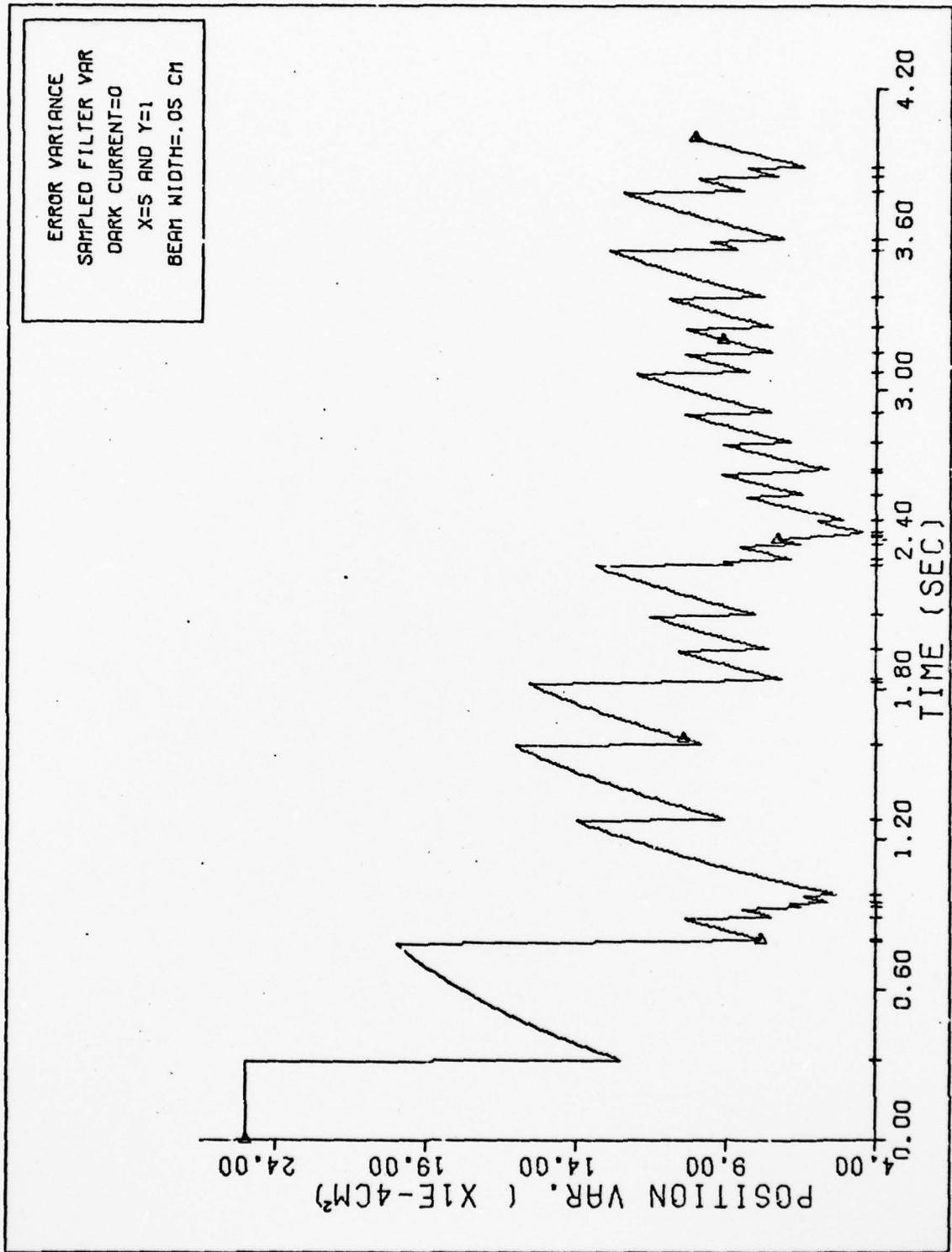
Figure 13b X-POSITION FILTER VARIANCE
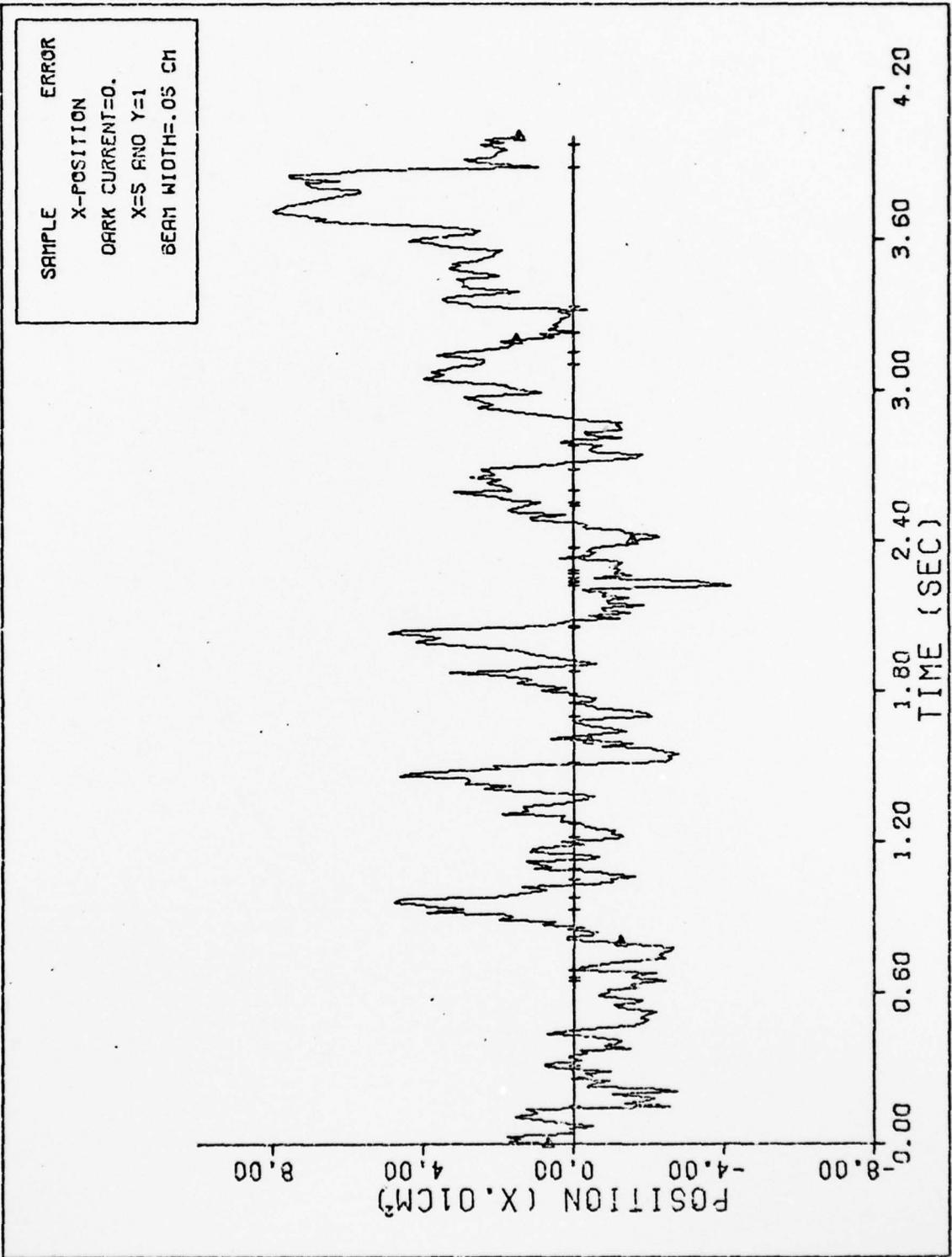
Figure 14a   X-POSITION

84
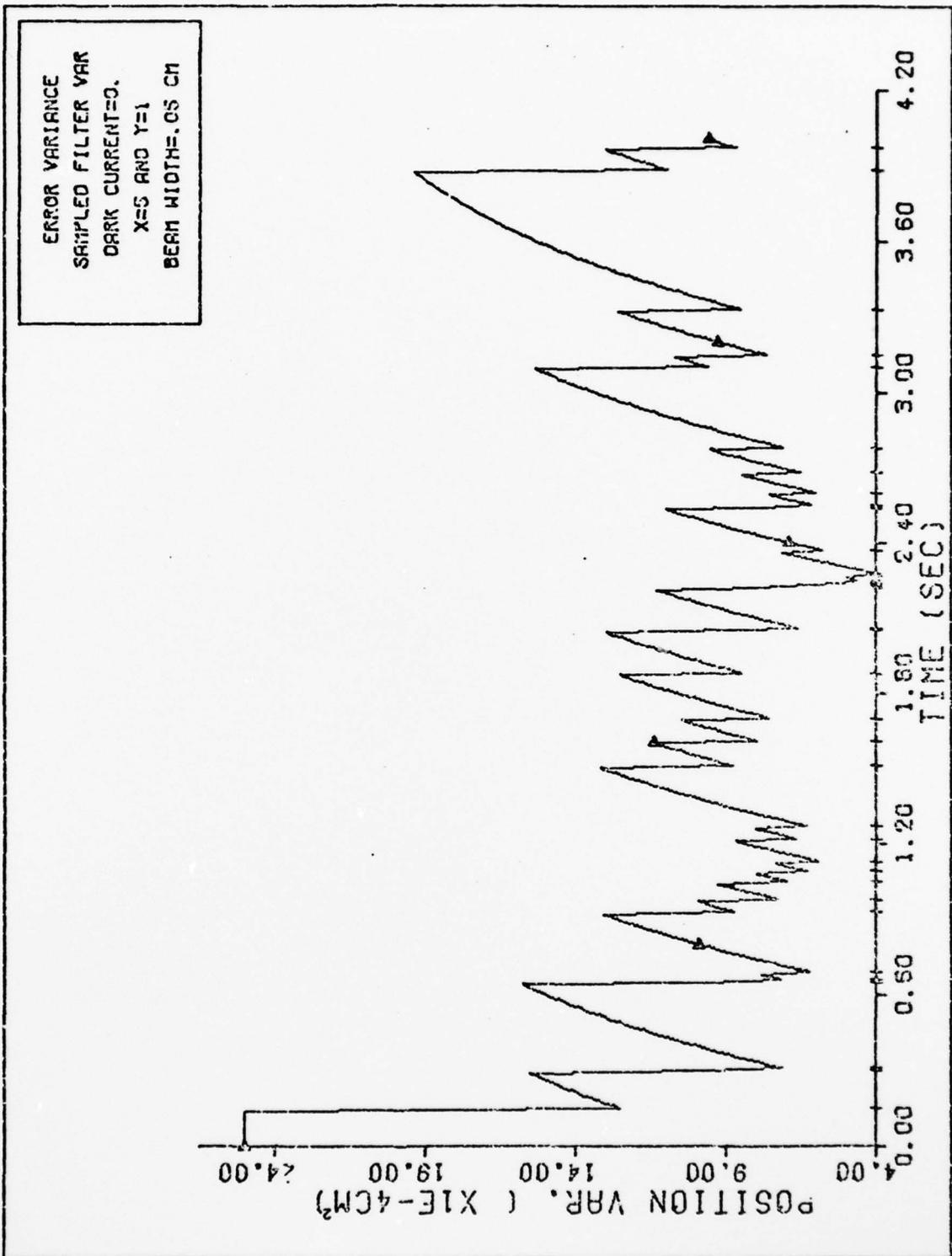
Figure 14b X-POSITION FILTER VARIANCE

85

a smaller filter growth in the variance, and the next measurement will have a lower weighting.

For this research, the coherence time, τ_c, of the spot is one second; for X=5, the term $\frac{nI_o}{nf_o}\sqrt{2\pi}\ \sigma$ in Eq. (63) is equal to 10 events per second. This is the average rate function of the Gaussian spot, after integrating $\lambda_g(t,r)$, defined in Eq. (14), over the detector length which approximates an infinite detector length with respect to the Gaussian beam. It is well known for a poisson process that the mean and variance of the counting process N(t), defined in Appendix A, are respectively $E[N(t)] = \iint_{\ell}\int_0^T \lambda(\alpha,\beta)d\alpha d\beta$ and $VAR[N(t)] = \iint_{\ell}\int_0^T \lambda(\alpha,\beta)d\alpha d\beta$. It then follows for T=4 that $\lambda_g(t,r)$ can be expressed as

$$\lambda_g(t,r) = \left[\frac{nI_o}{nf_o}\sqrt{2\pi}\ \sigma\right]\left[\frac{1}{\sqrt{2\pi}\ \sigma}\ \exp(\frac{-(r-\underset{\sim}{x}(t))^2}{2\sigma^2})\right]$$

$$= \frac{10}{\sqrt{2\pi}\ \sigma}\ \exp(\frac{-(r-\underset{\sim}{x}(t))^2}{2\sigma^2}) \tag{64}$$

so that E[N(t)]=40 events with standard deviation of 6.3 events. In Figs. (13) and (14), the number of events for this particular case are 34 and 36 events, respectively. Also, in these sample runs, simultaneous events did not occur. To the second moment level, these results indicate some validity in the Binomial Method in generating the point process observations in time and space.

Finally, the test cases for this method are plotted in Figs. (15) through (17); note the different scales on each plot. As expected, increasing Y (the spot's jitter normalized

6473

1.0

4.5
5.0
5.6

2.8

2.5

3.2

2.2

3.6

2.0

4.0

1.1

1.8

1.25

1.4

1.6
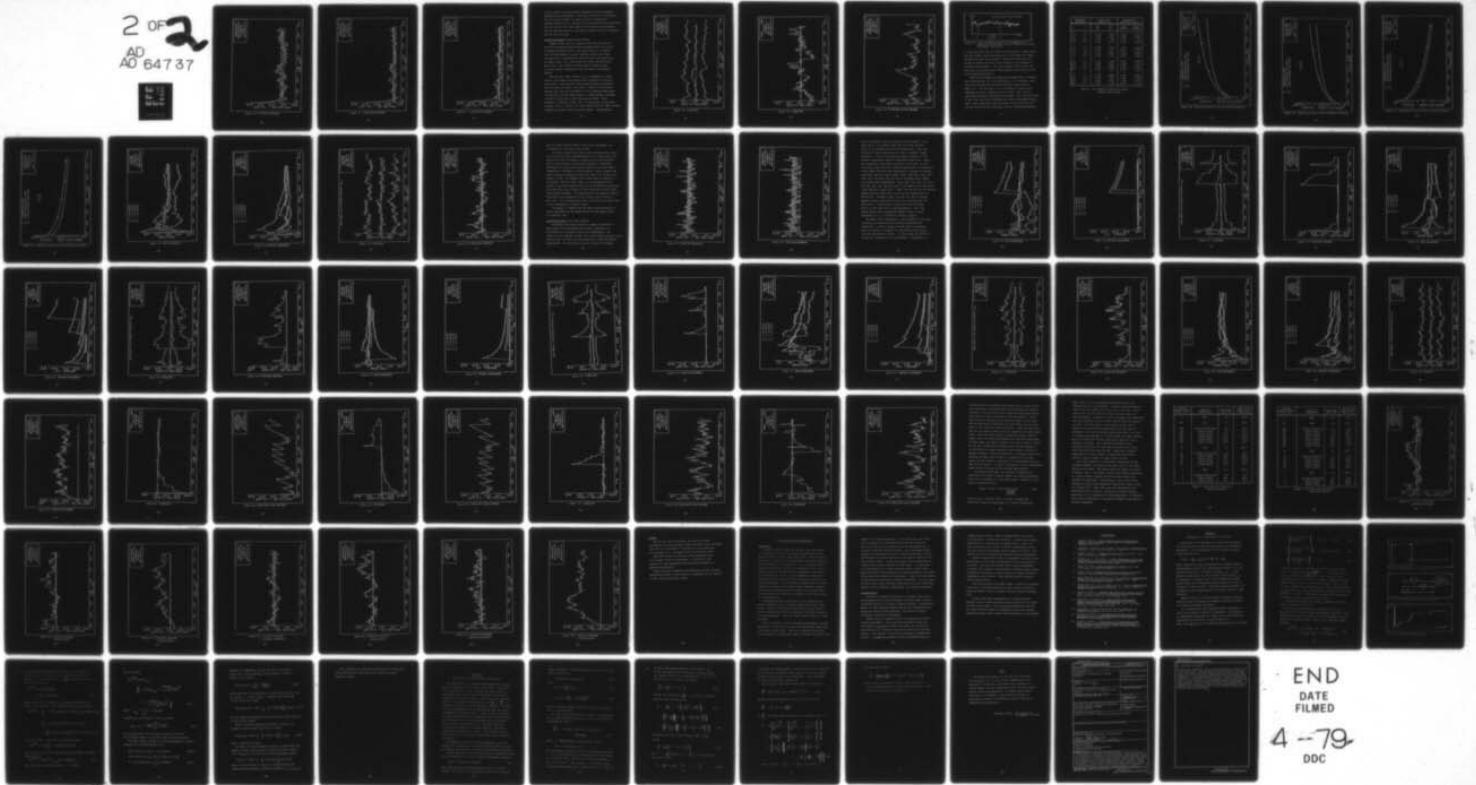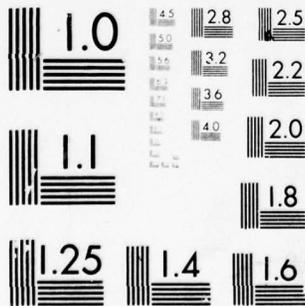
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A
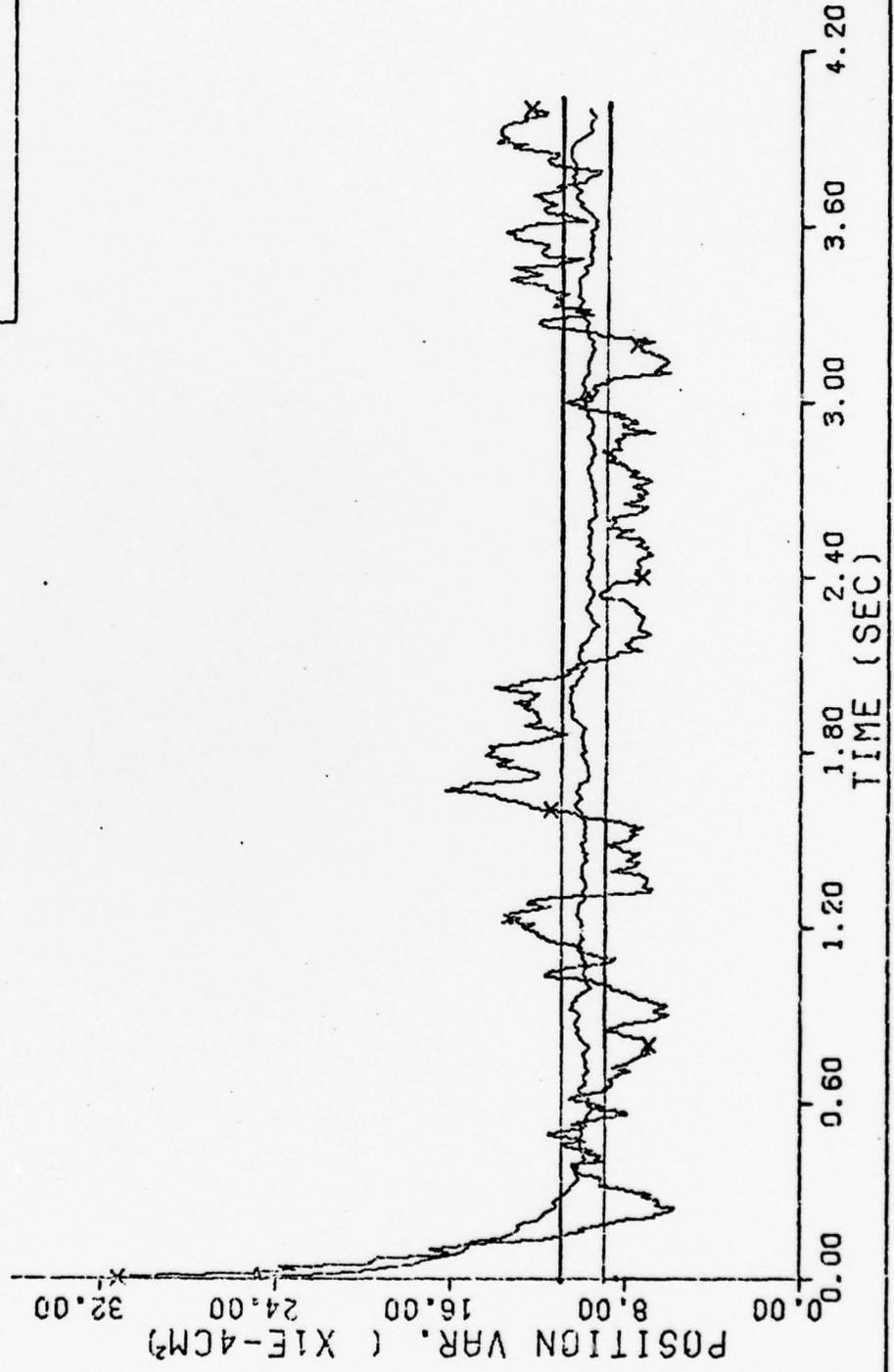
Figure 15   X-POSITION VARIANCES

87
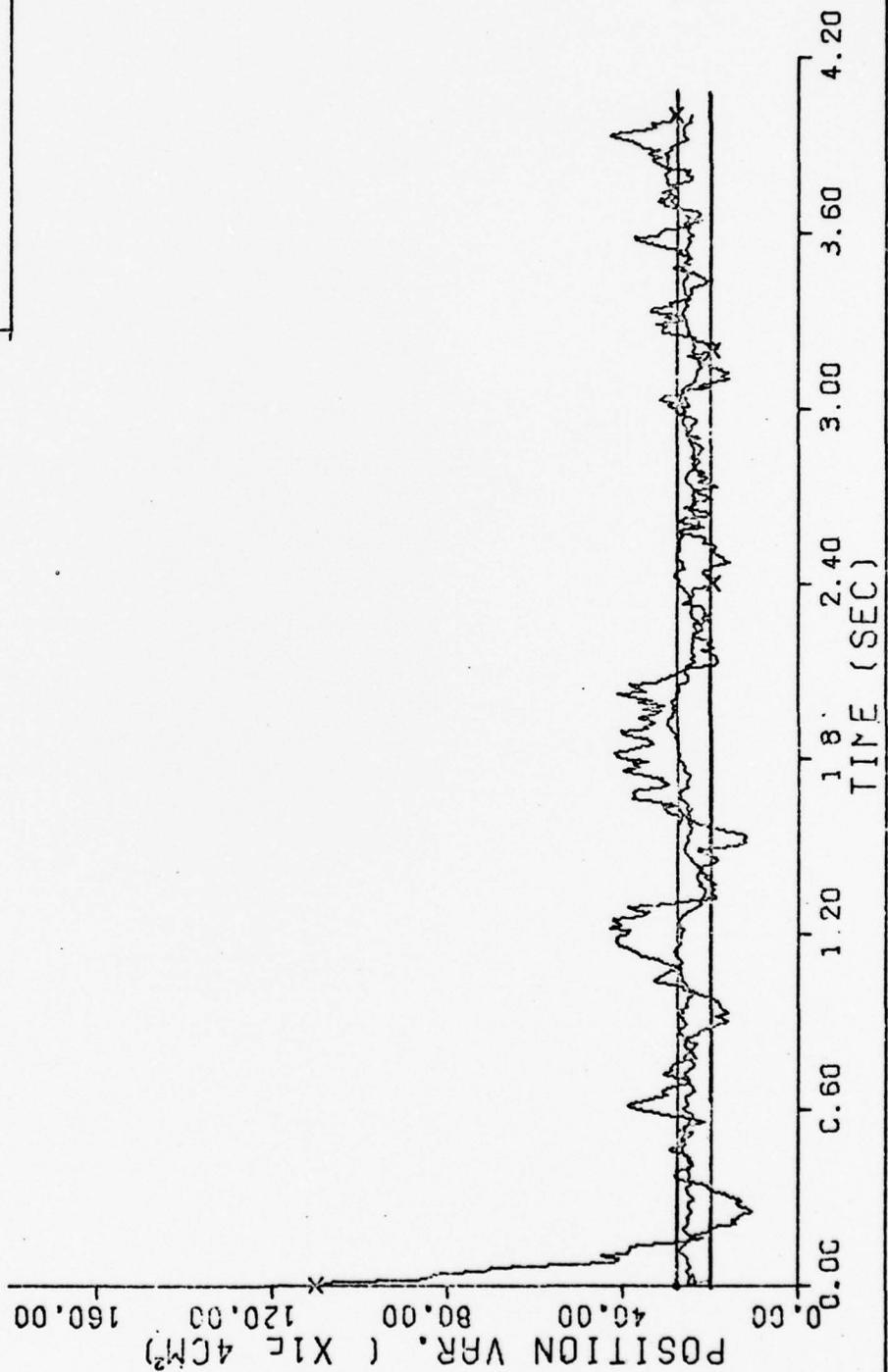
Figure 16   X-POSITION VARIANCES
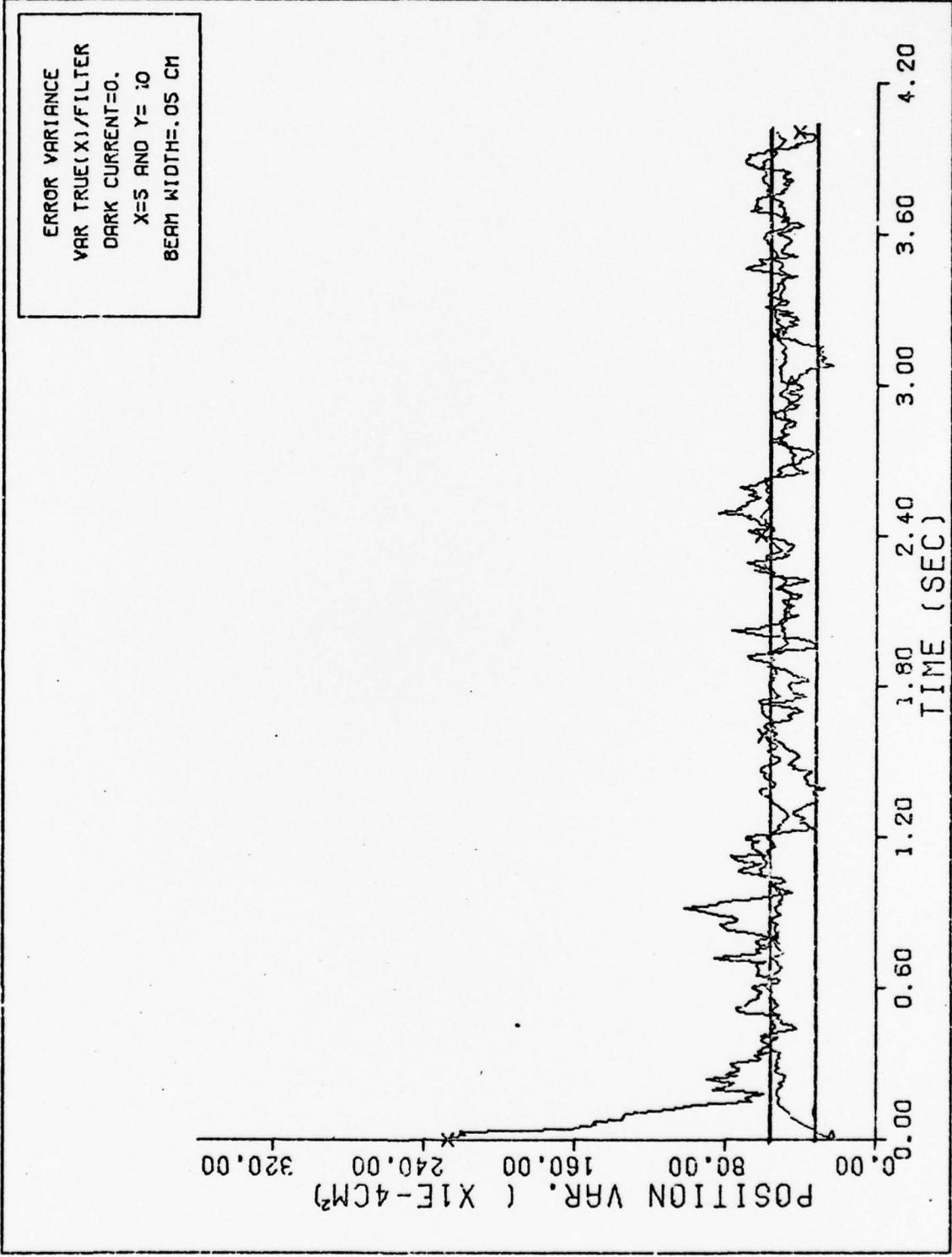
Figure 17 X-POSITION VARIANCES

89

by the square of the beam width) increases the error variances –
implying poorer performance. Even when the filter is initialized
with its state estimate not equal to the true position of
maximum intensity (i.e., initialize from non-zero initial conditions),
the plot in Fig. (18) shows that the error process is zero-mean.
This is expected since the conditional estimate of $\underset{\sim}{x}(t)$ is unbiased
and $\underset{\sim}{x}(t)$ has zero mean.

Conditional Method (without dark current)

Shown in Fig. (19) is a sample of the error process and
the filter variance generated by the Snyder Filter for the set
of parameters, X=20 and Y=1. From these plots, it is evident
that the measurements of event locations at the event times
are important in minimizing the filter's error and uncertainty.
As shown in Fig. (19a), the errors are within one beam width;
and in Fig. (19b), there is little filter variance growth
as events appear frequently in time (as seen in the Binomial
Method).

For the test cases in Table 2, it is desirable to choose
values of the sample and average filter variances at selected
time points. For each case, these variance points are plotted
with the upper and lower bound curves. However, it was noted
earlier that there are fluctuations in the variances which make
it difficult to select these points. A "time-average" technique
is used to select and to plot these points easily. This
technique is depicted in Fig. (20). In this case, a fixed time
interval, that is chosen to be much greater than the sample period,
contains time points that are equally spaced. Corresponding
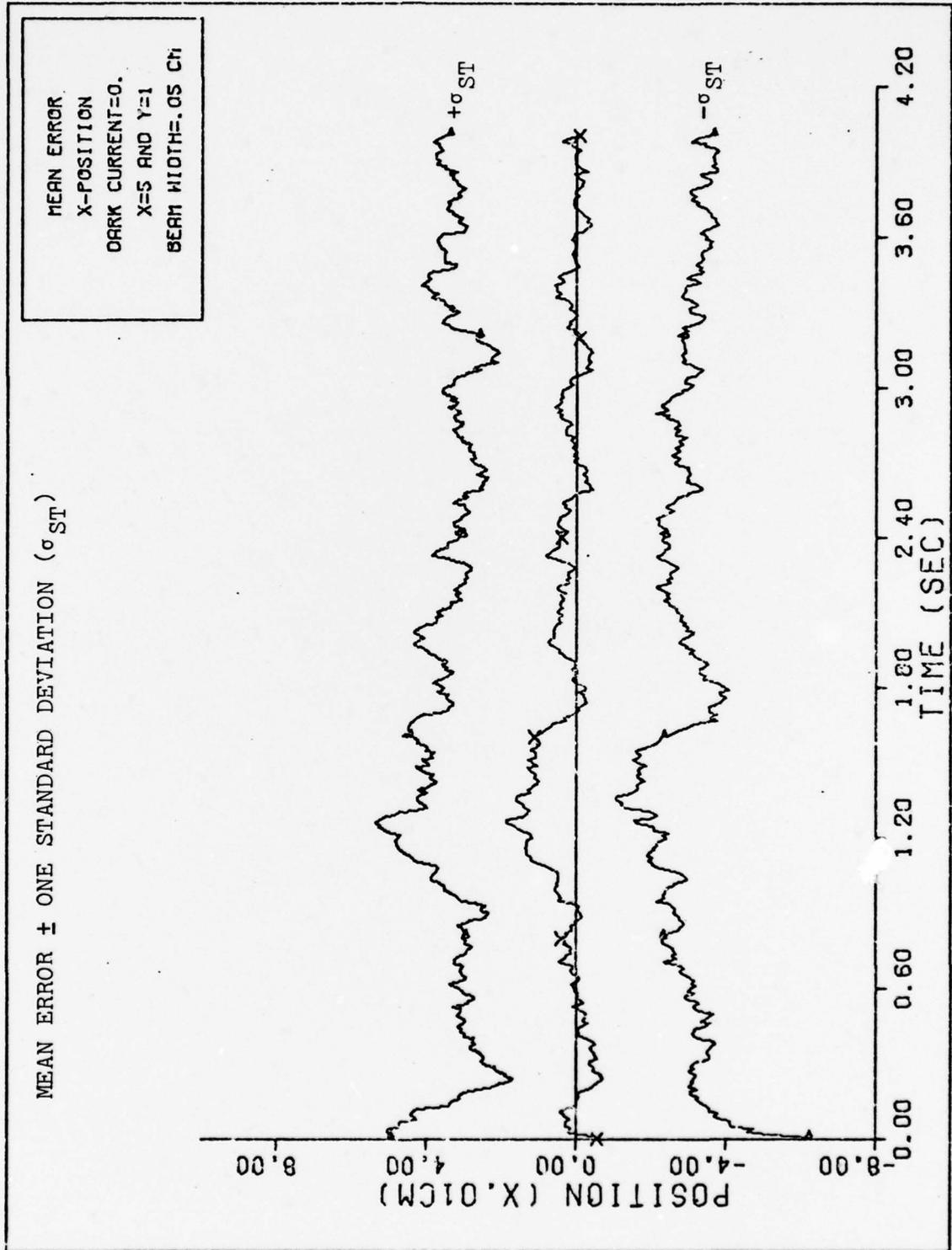
Figure 18   X-POSITION

91
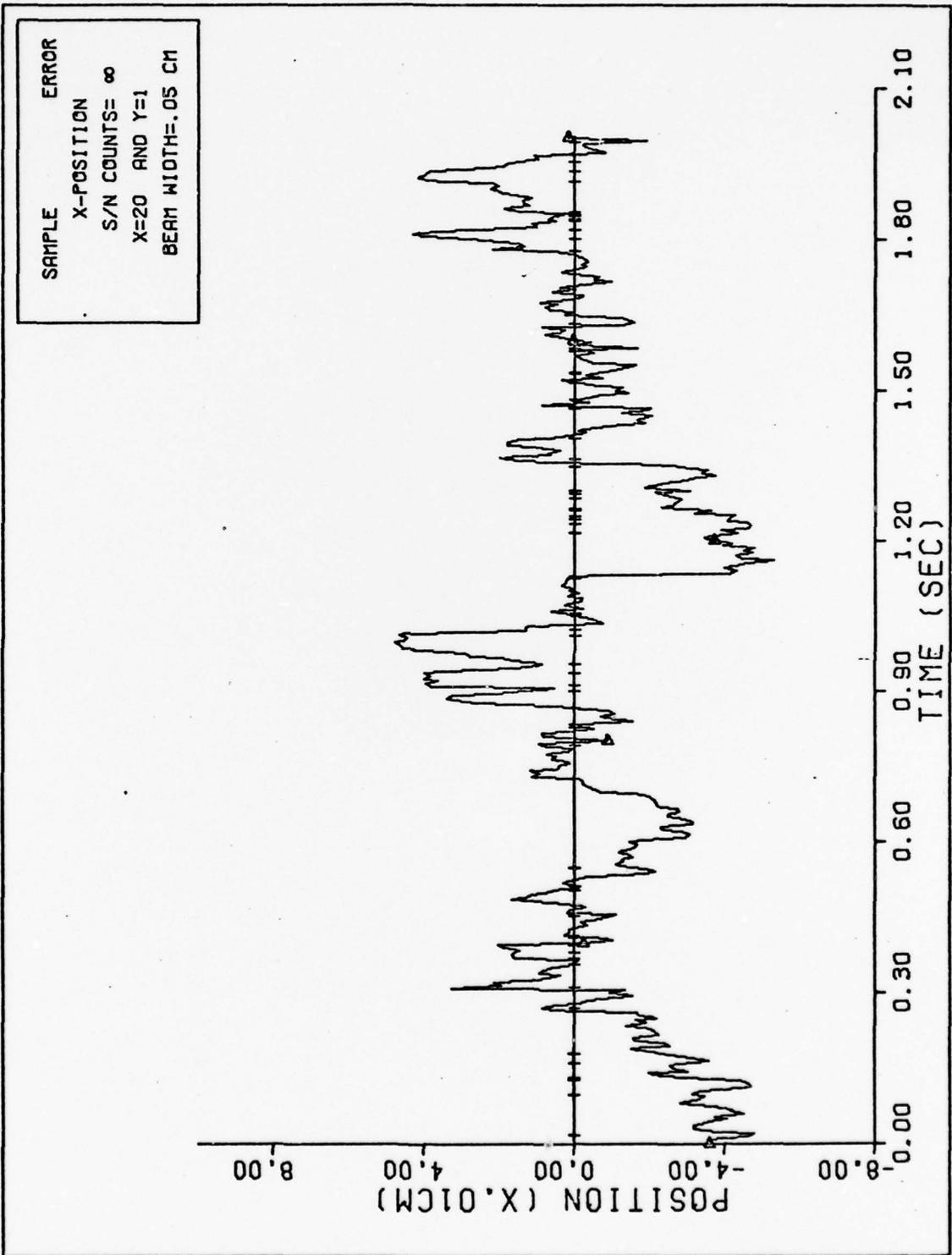
Figure 19a X-POSITION
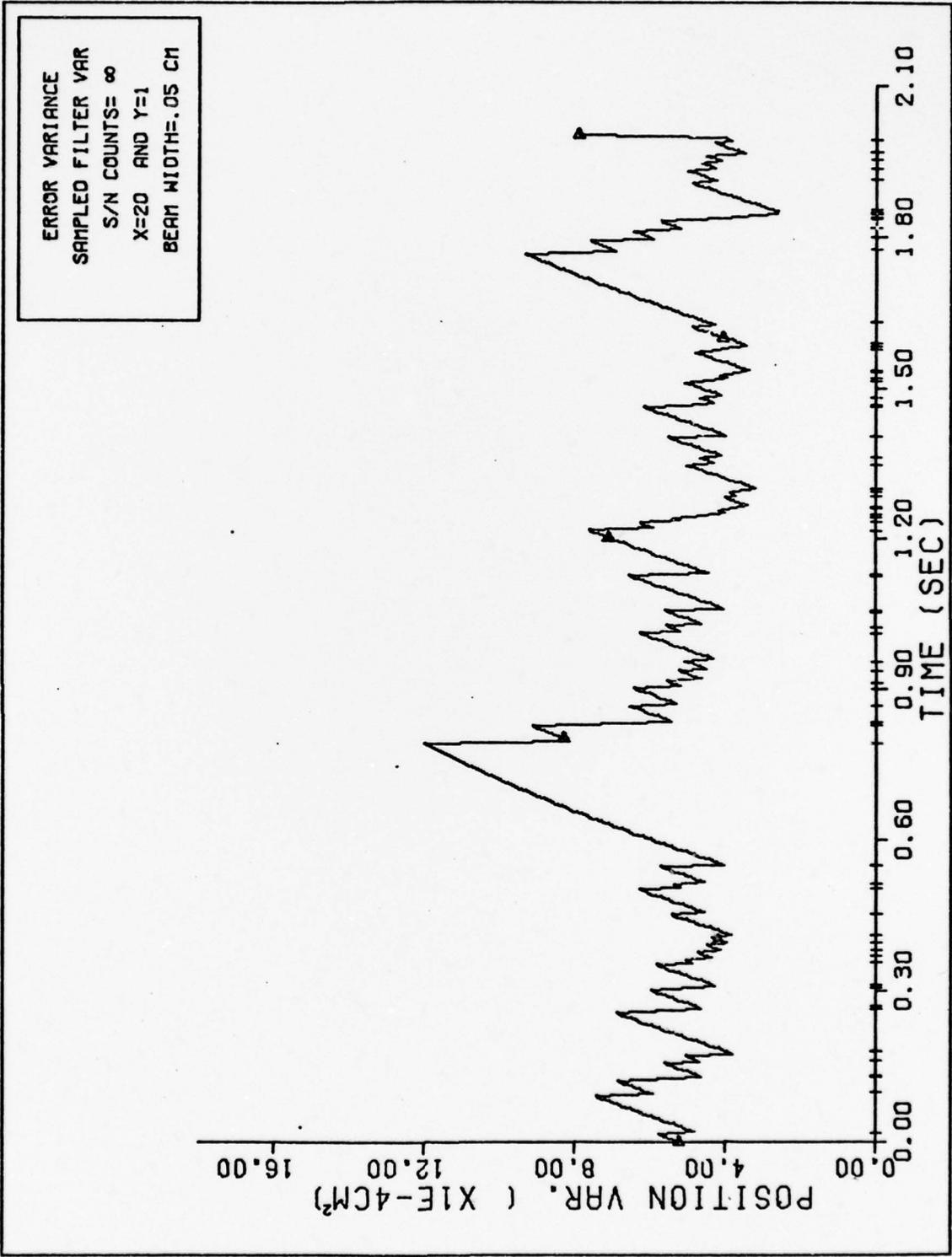
Figure 19b X-POSITION FILTER VARIANCE
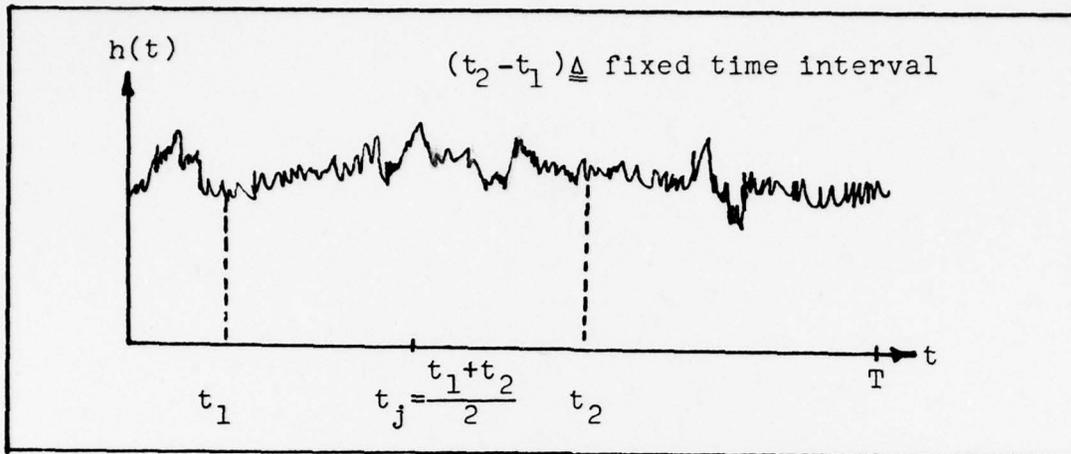
Figure 20   "Time-average" Technique (h(t) denotes time
            history of either true or filter variance calculated
            from sample statistics)

to each time point are the values of the variances. These values
are then summed, and the resulting sum is divided by the number
of points in that fixed interval. This final value will then
be used to denote the value of the variance at the midpoint of
the fixed time interval. This time interval "runs" across the
time axis to calculate these "time-average" values until the end
of the observation time, T.

With this technique, the sample and average filter variances
are listed in Table 3 at the time point, t_j = 1.5 second. These
values are plotted with the upper and lower bound curves in
Figs. (21). The time point, t_j = 1.5 second, is chosen after
examining the plots, similar to plots in Fig. (22), for each
case which indicates steady-state is reached. Note that the values
of the true and filter variances are close to the upper bounds,
which seems reasonable, since the values of X=1, X=5, Y=1, and
Y=5 are relatively poor conditions for good tracking performance.

94

Performance Parameter		Normalized Bounds on MSE		Normalized Simulation Values	
X	Y	$\dfrac{P_*}{\sigma^2}$	$\dfrac{P^*}{\sigma^2}$	True Variance	Filter Variance
5.0	0.5	0.232	0.250	0.233	0.251
5.0	1.0	0.358	0.408	0.388	0.406
5.0	5.0	0.905	1.305	1.340	1.270
5.0	10.0	1.320	2.240	2.370	2.180
1.0	0.5	0.366	0.390	0.340	0.389
1.0	1.0	0.618	0.707	0.621	0.703
1.0	5.0	1.790	2.870	2.570	2.845
1.0	10.0	2.900	5.420	4.900	5.380
0.5	5.0	2.320	3.590	4.200	3.490
10.0	5.0	0.659	0.880	0.935	0.854
15.0	5.0	0.545	0.698	0.670	0.654
20.0	5.0	0.496	0.542	0.592	0.558
0.5	1.0	0.732	0.816	0.898	0.803
10.0	1.0	0.270	0.302	0.319	0.298
15.0	1.0	0.227	0.250	0.243	0.242
20.0	1.0	0.200	0.218	0.222	0.212

Table 3. Test Cases for Conditional Method
(without dark current)
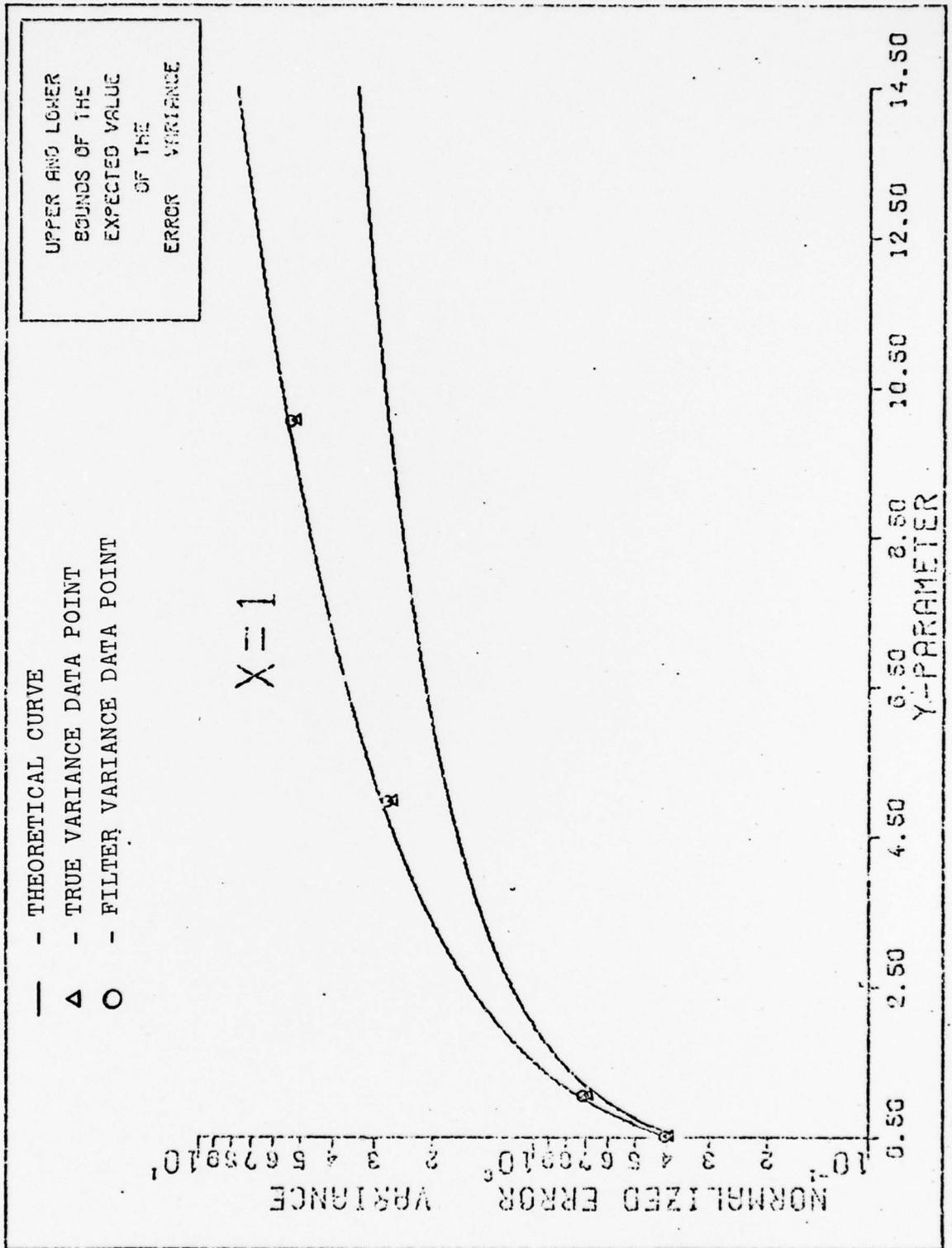
Figure 21a   Simulation Values of True and Filter Variances
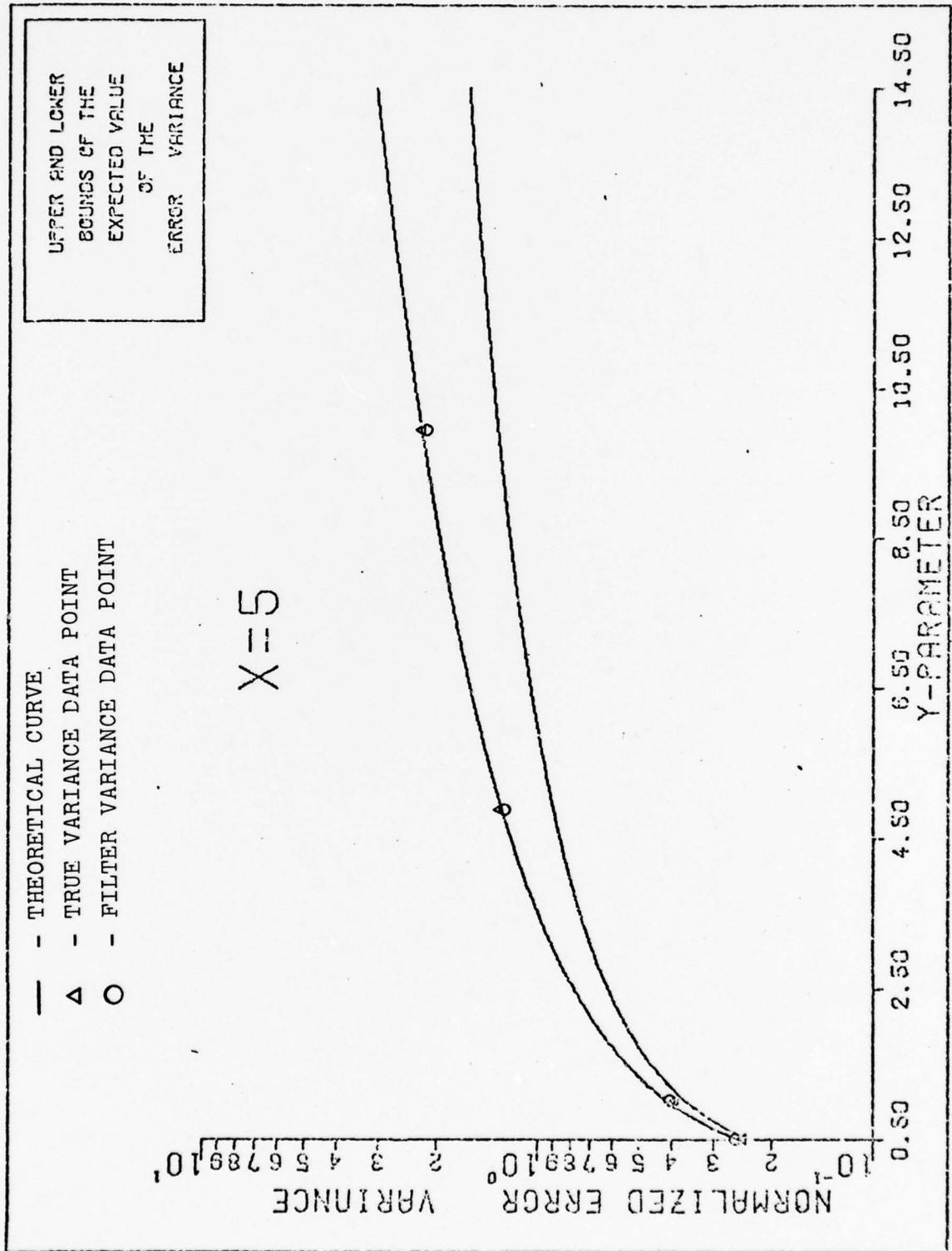
Figure 21b   Simulation Values of True and Filter Variances
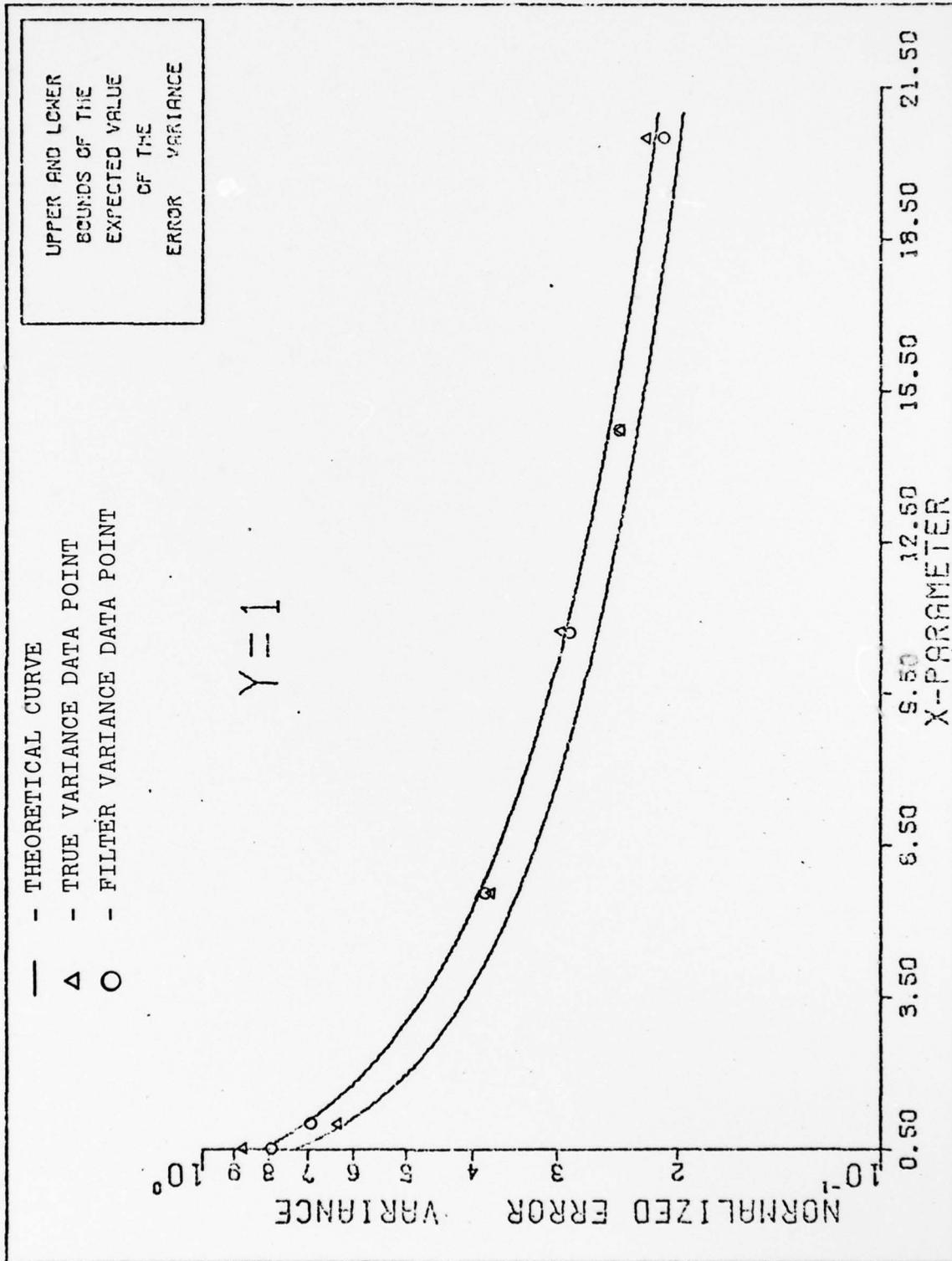
Figure 21c   Simulation Values of True and Filter Variances
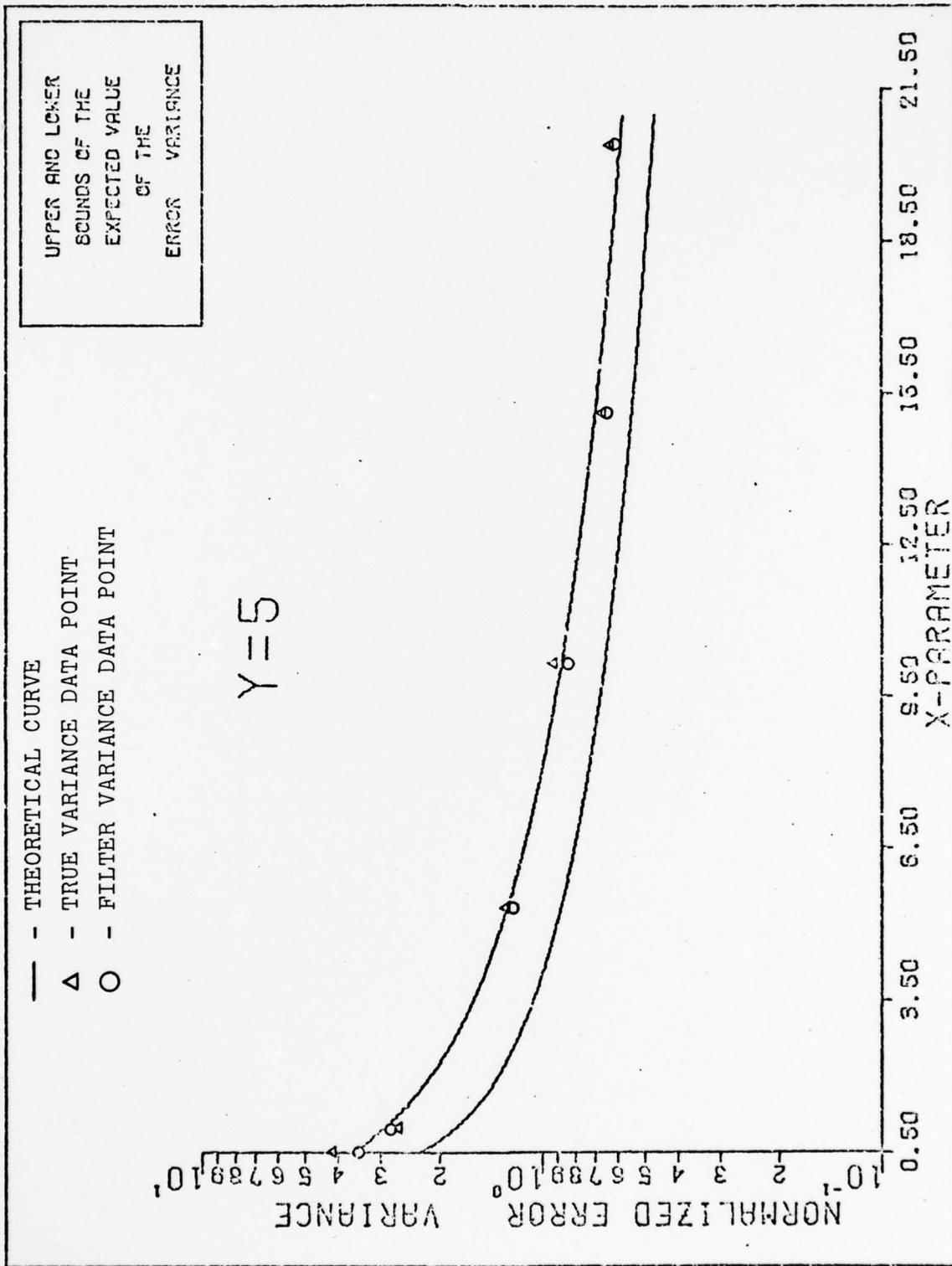
Figure 21d   Simulation Values of True and Filter Variances
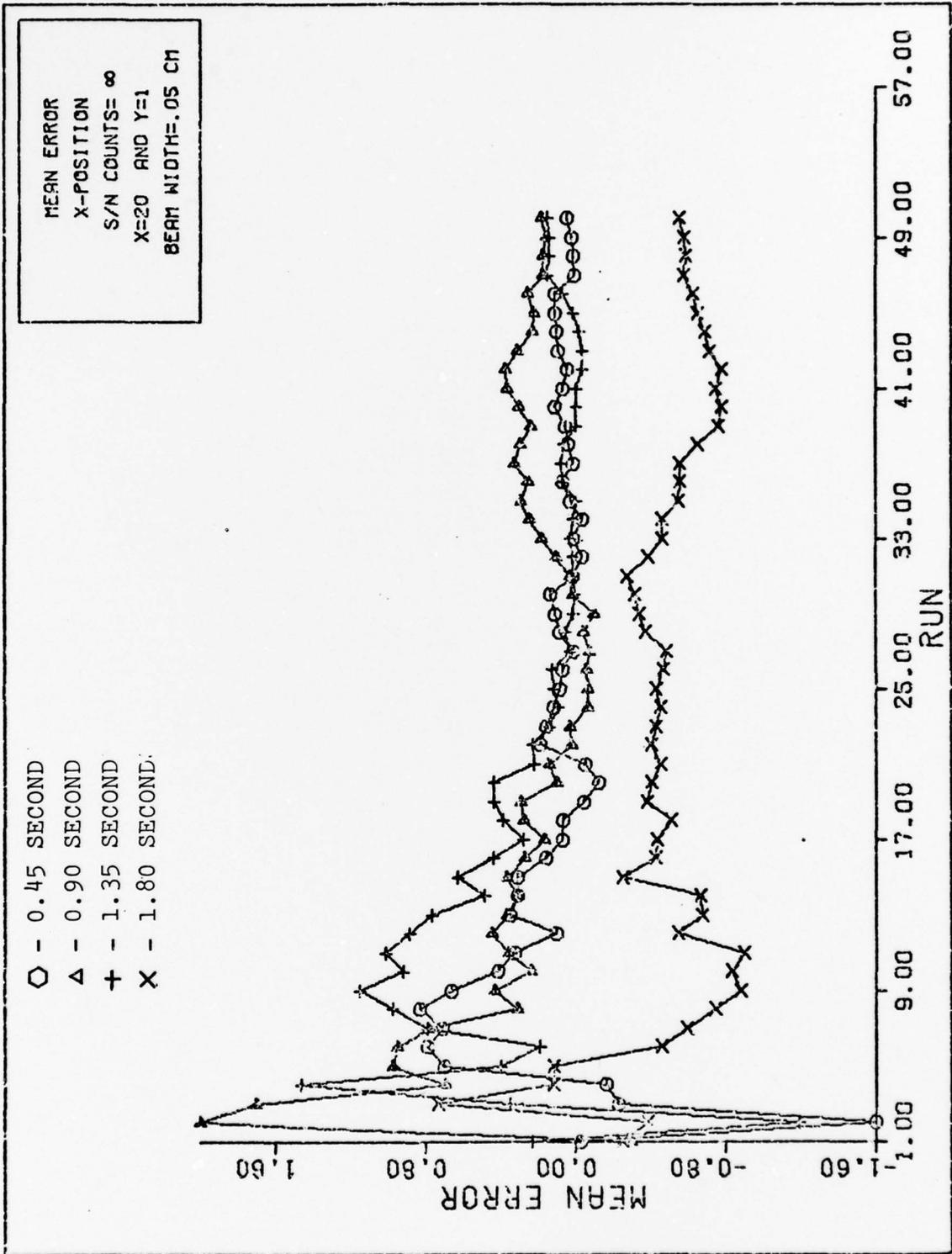
Figure 22a MEAN CONVERGENCE
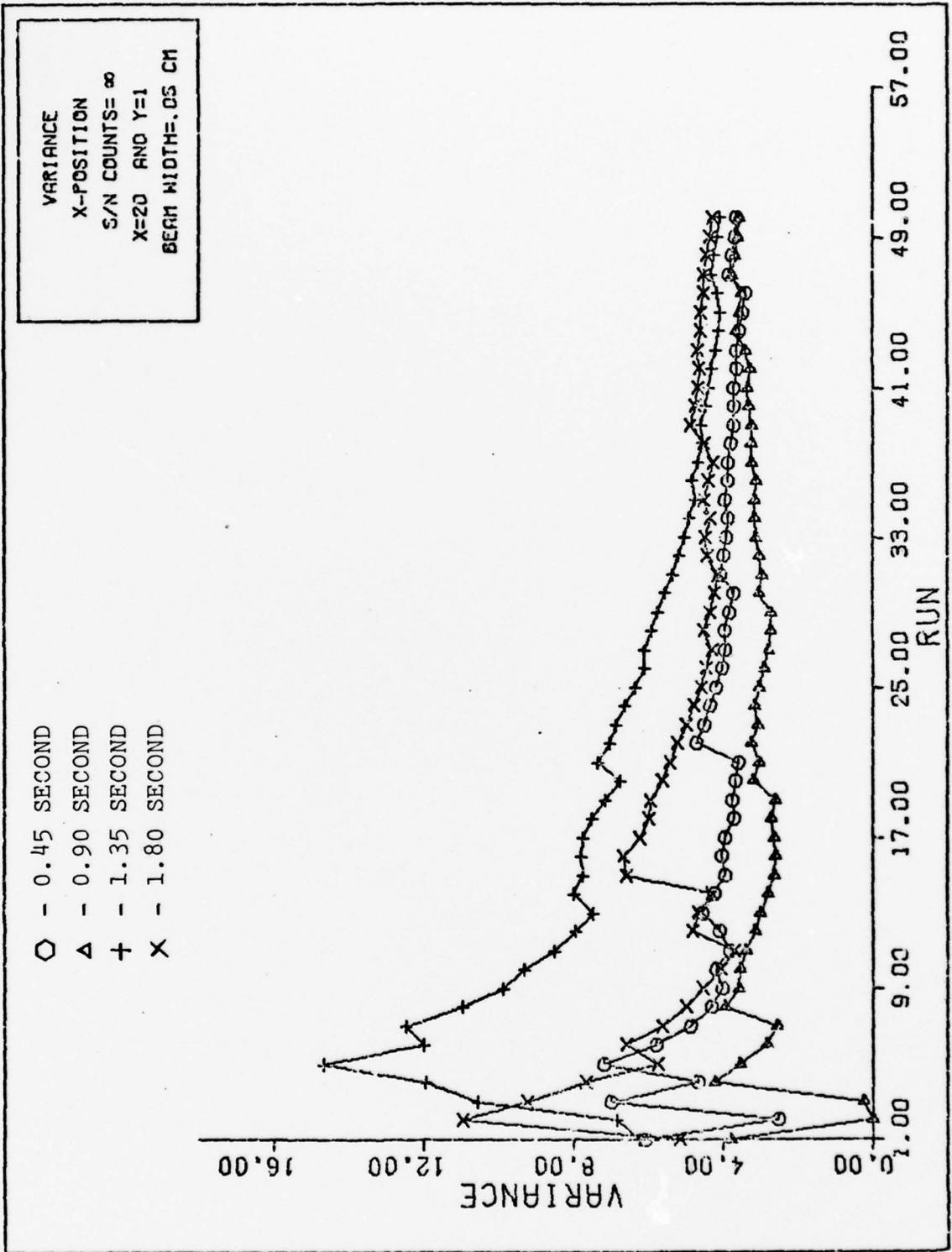
Figure 22b VARIANCE CONVERGENCE

101
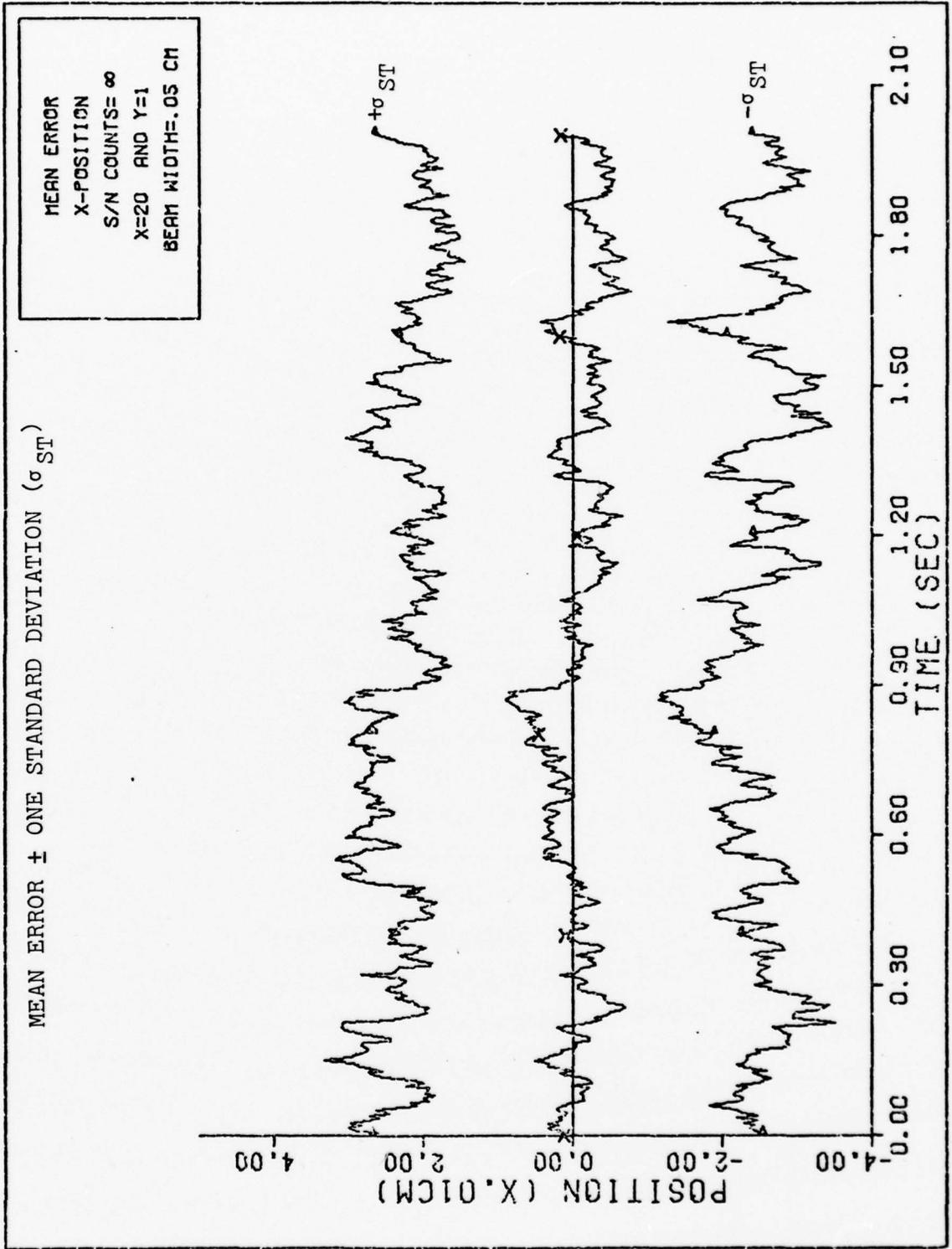
Figure 22c X-POSITION

102
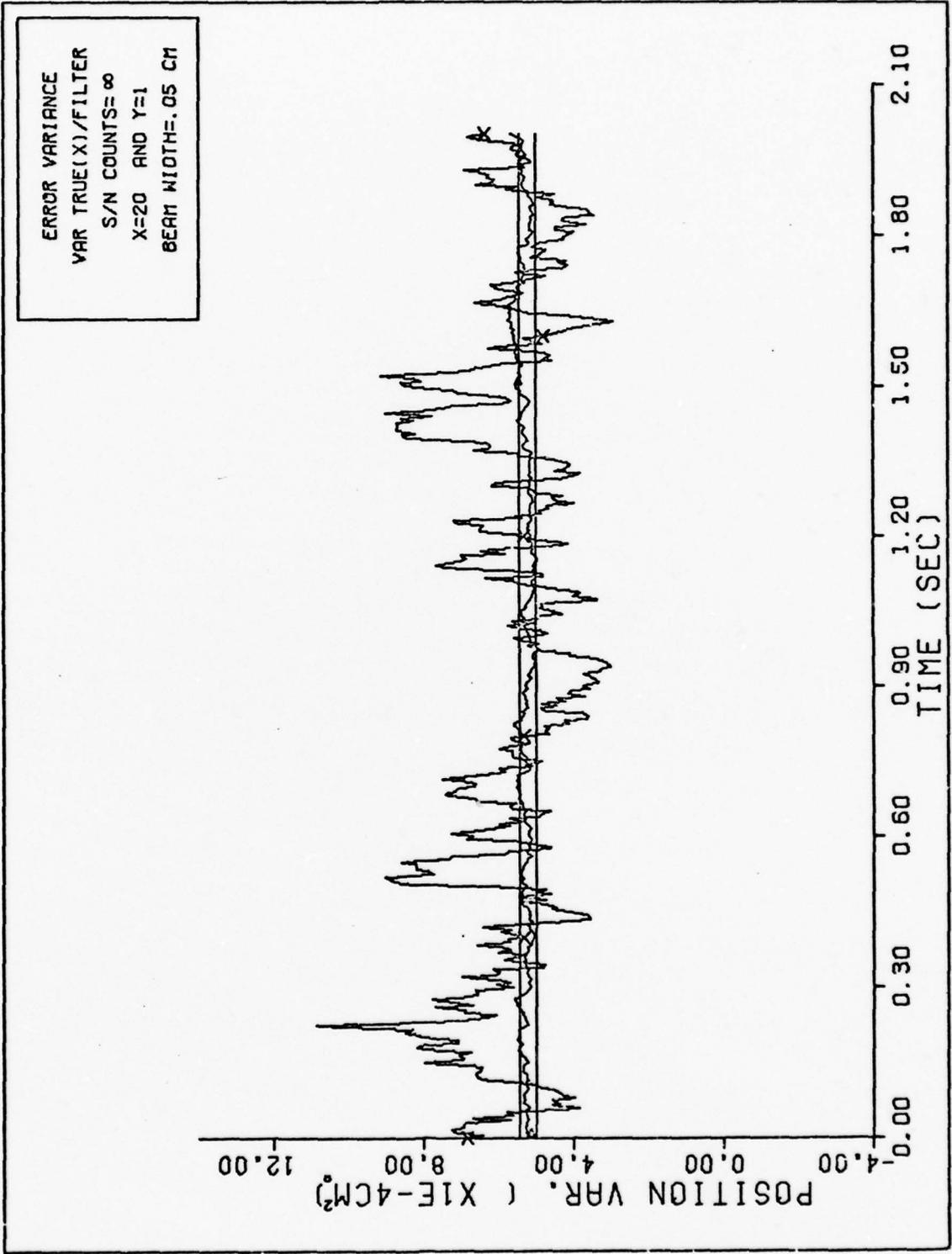
Figure 22d X-POSITION VARIANCES

103

That is, these results indicate that actual performance can be described by the upper bound on MSE.

Two issues should be of concern when plotting these values. One is that the filter performance must be in steady state, and the other is that the number of sample runs in the simulation must be large enough to yield some confidence of inferring information from the error variance generated by either sample statistics or averaging of filter outputs. Hence, a particular case is chosen to extend the simulation time to 40 seconds; the results in Fig. (23) indicate that steady state is reached in one second. Then the issue again is of confidence in the values of the true variances which can be increased by performing more simulation runs; because of time constraints, fifty simulations remain unchanged. To illustrate that the number of simulation runs is not appropriate in all cases, refer to plots in Fig. (22). For this particular case, X=20 and Y=1, the mean error is not converging to a definite value.

In summary, it appears that the upper bound describes the actual performance of the Snyder Filter for few signal counts per coherence time.

Conditional Method (with dark current)

Particular cases were selected to examine the effects of dark current and to determine the filter's sensitivity in performance as it processes these noise measurements. The results of this analysis is that, for average signal-to-noise counts ratio of 500, the filter was very sensitive to dark distributed measurements. The results are plotted in Figs. (24) through

104
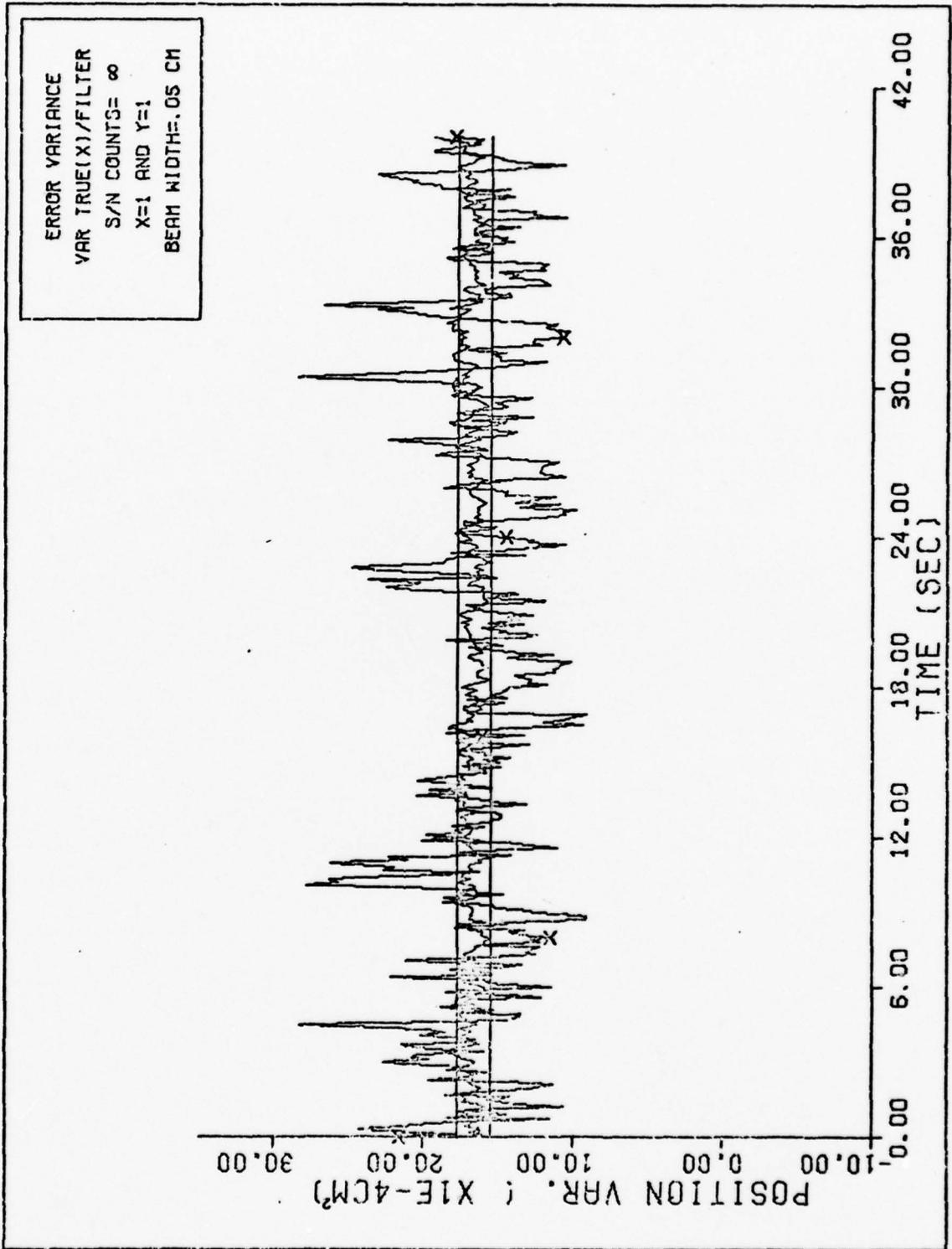
Figure 23a X-POSITION VARIANCES

105
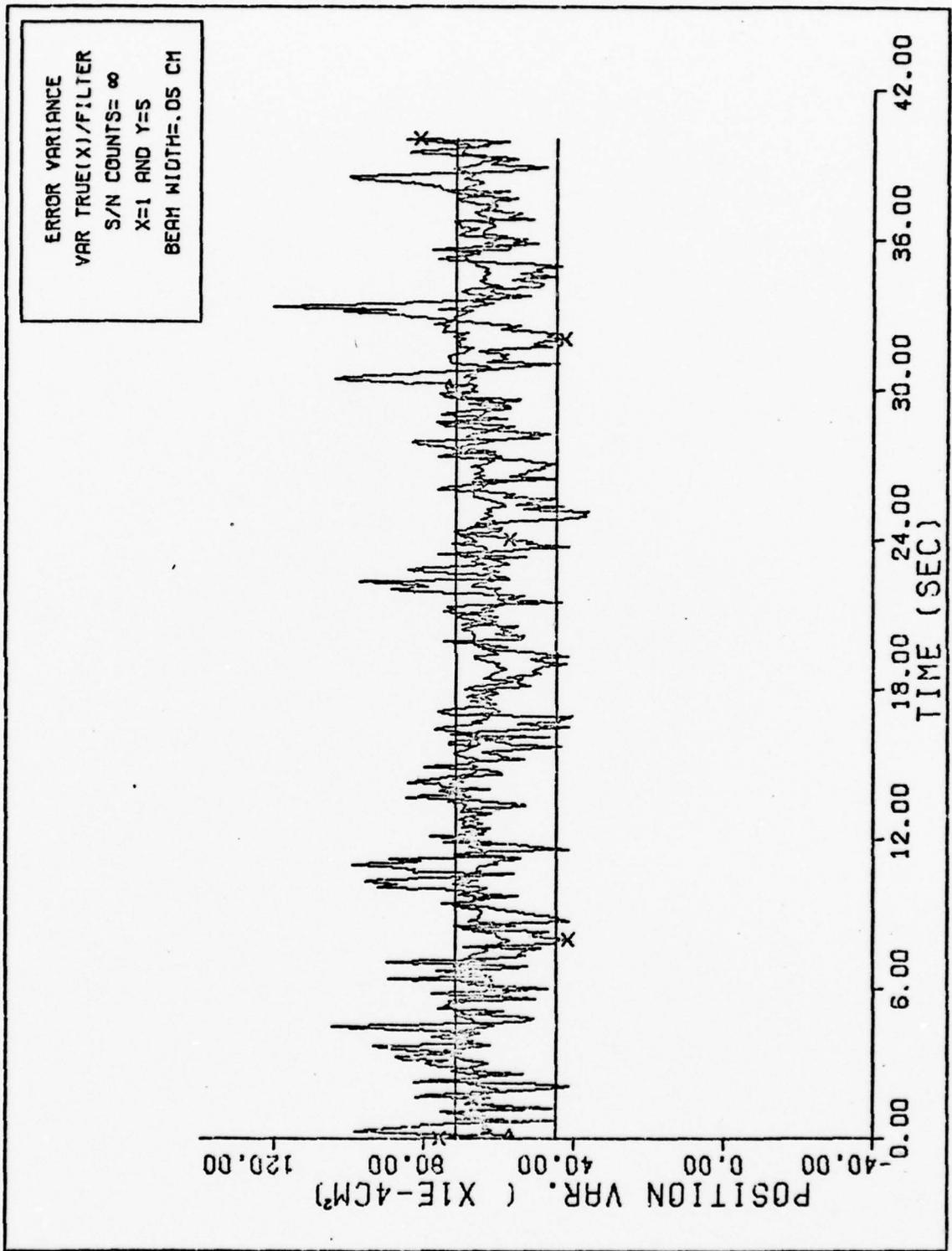
Figure 23b X-POSITION VARIANCES

106

(28) for different cases, as indicated by the figures. From these plots, it is clearly shown that the filter performs poorly if it processes even one dark measurement. Note that the plots in these figures are not close to ensemble average statistics. That is, to obtain the sample variance, it will require a substantial amount of additional sample runs. Again, the confidence in the sample statistics is an issue. On the other hand, these results do provide insight into the sensitivity of the Synder Filter to dark measurements. However, to evaluate individual effects of dark current, Figs. (24a, 24b), (25a, 25b), (26a, 26b), and (28a, 28b) are used to determine which sample runs contain dark measurements. Thus, Figs. (29a, 30a, 31a, 32a) and (29b, 30b, 31b, 32b) are plots of the sample errors and sample filter variance, respectively. Figs. (29) and (30) are for average signal-to-noise counts ratio of 500 and 50, respectively, for X=5 and Y=1. Similarly, Figs. (31) and (32) are for average signal-to-noise counts ratio of 500 and 5, respectively, for X=20 and Y=1. These results clearly show that the Snyder Filter is very sensitive to dark measurements. Thus, the Snyder Filter cannot distinguish between a dark or signal measurement (i.e., it treats all events as signal).

To enhance filter performance, filter-tuning (in this case, increasing the beam-width) and residual monitoring were considered. In filter tuning, the beam width is increased, which corresponds to weighing the measurements less heavily from the photodetector surface. In other words, the filter is having less confidence in its measurements. Conversely, if
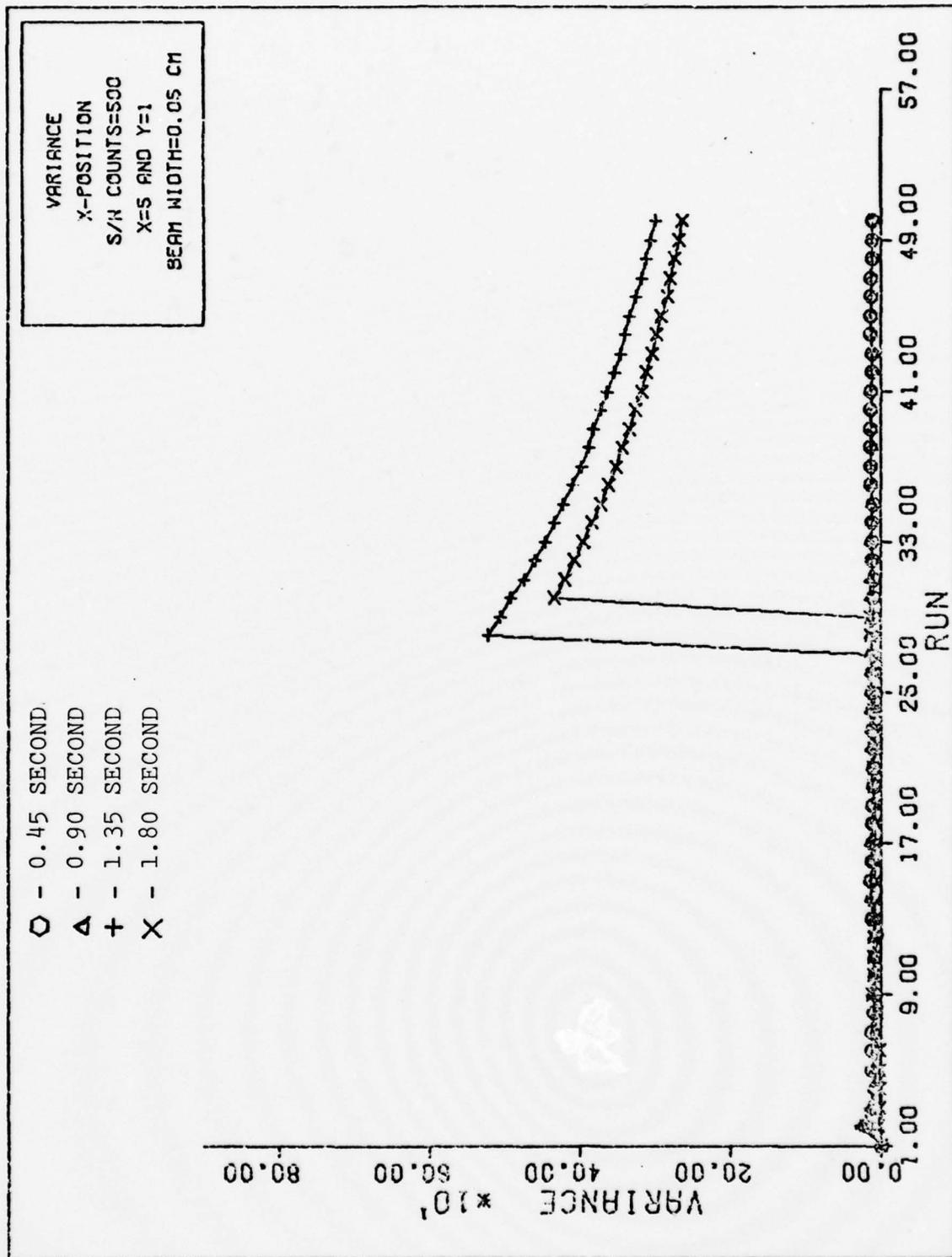
107
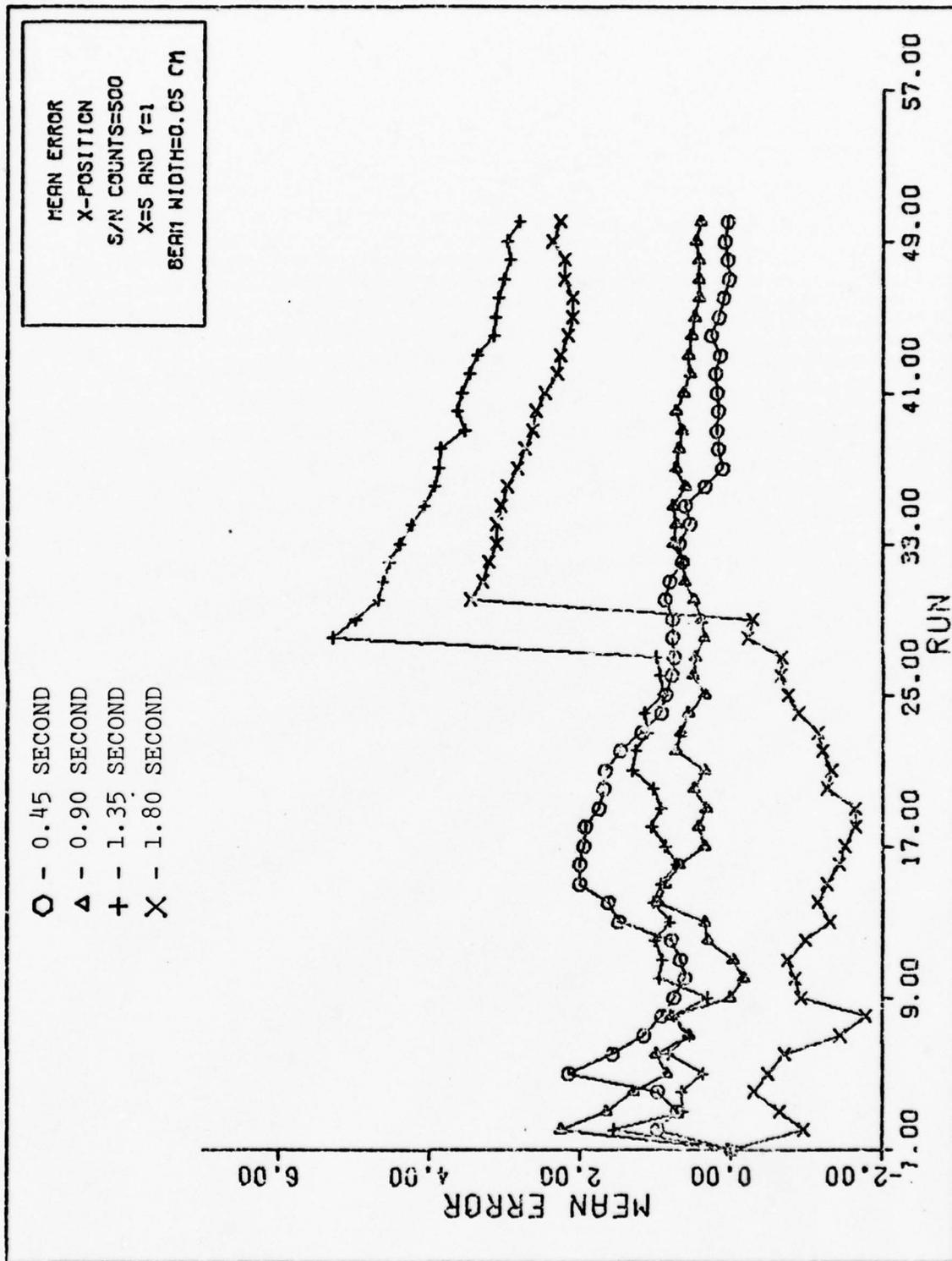
Figure 24a MEAN CONVERGENCE
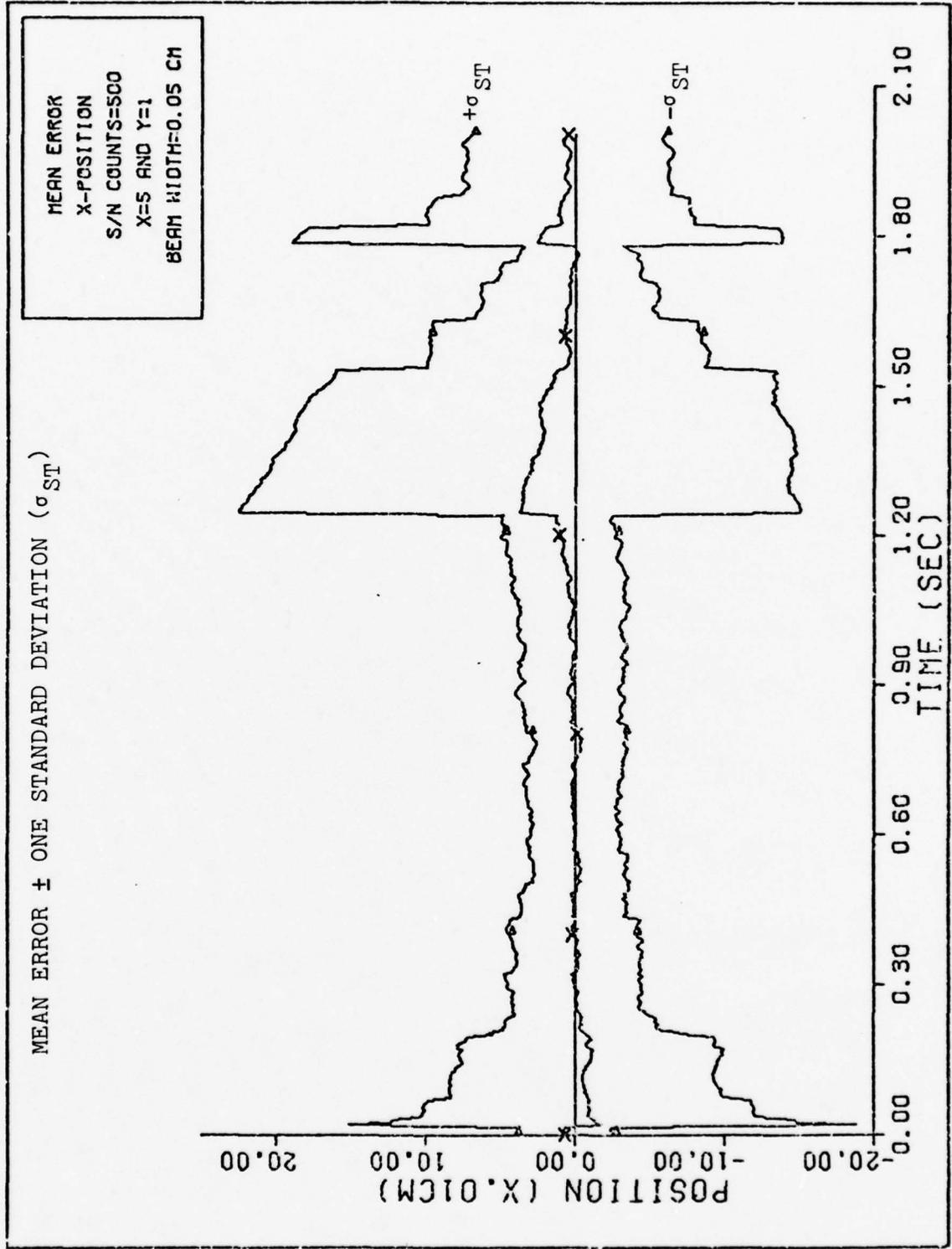
108

Figure 24b VARIANCE CONVERGENCE
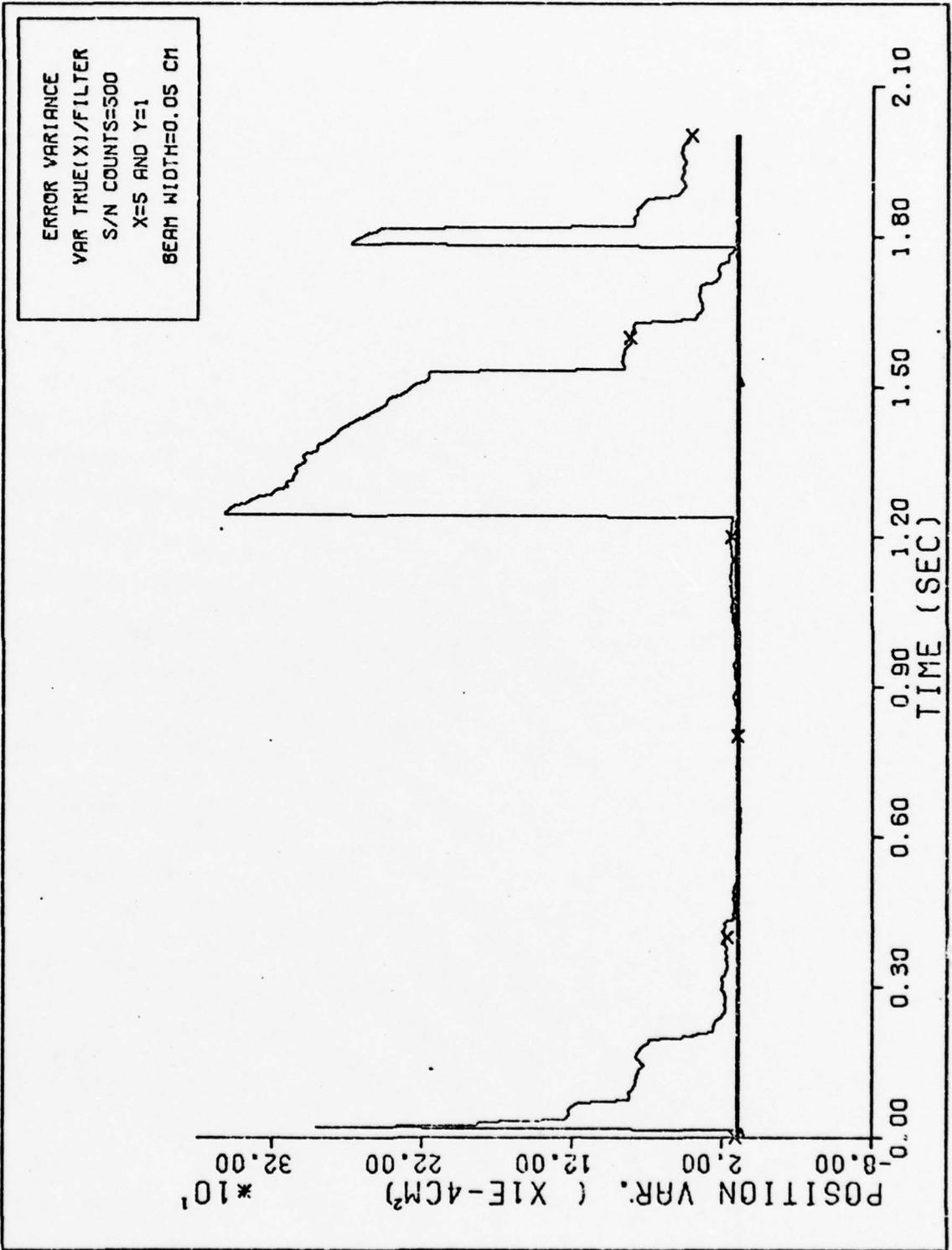
109

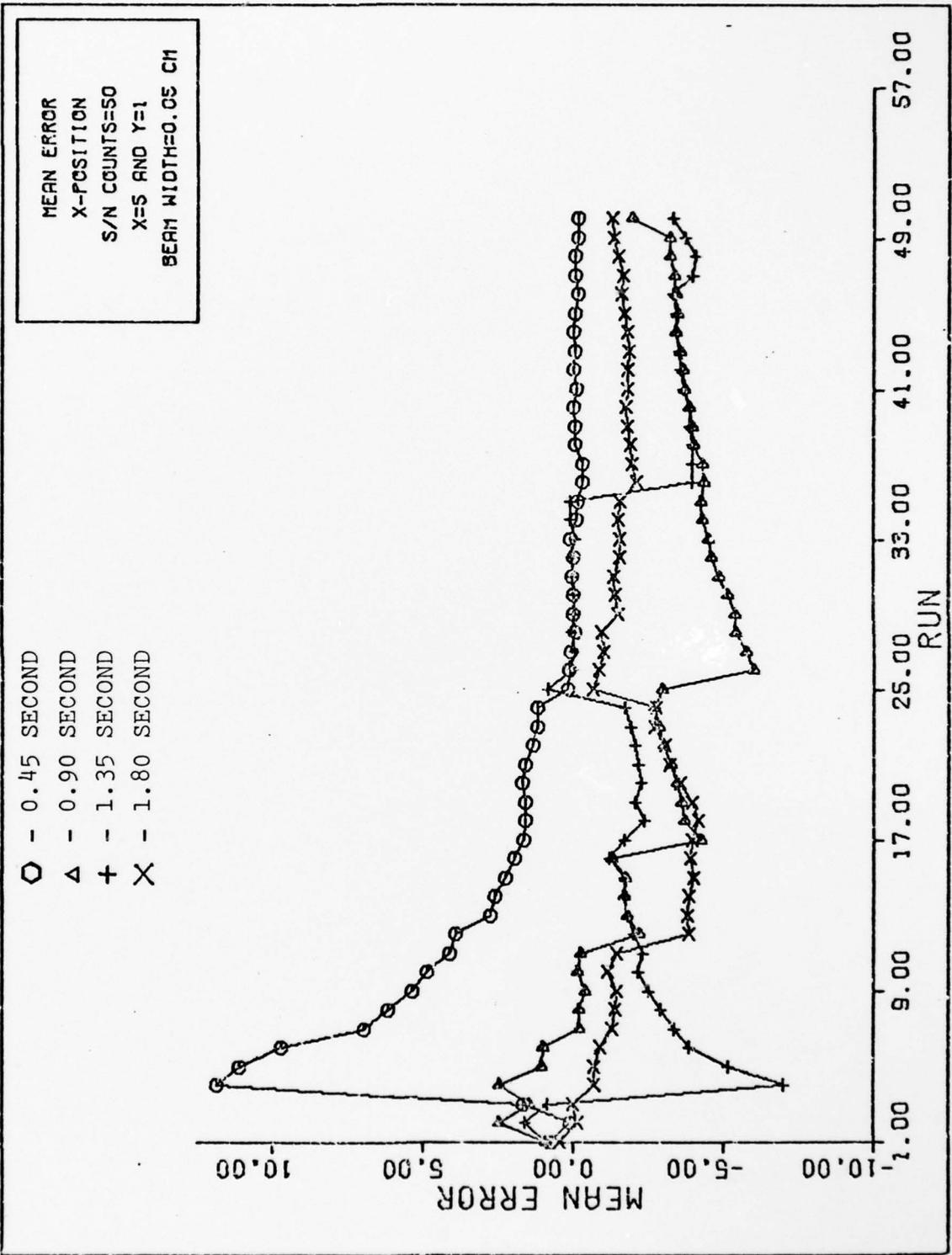Figure 24c X-POSITION

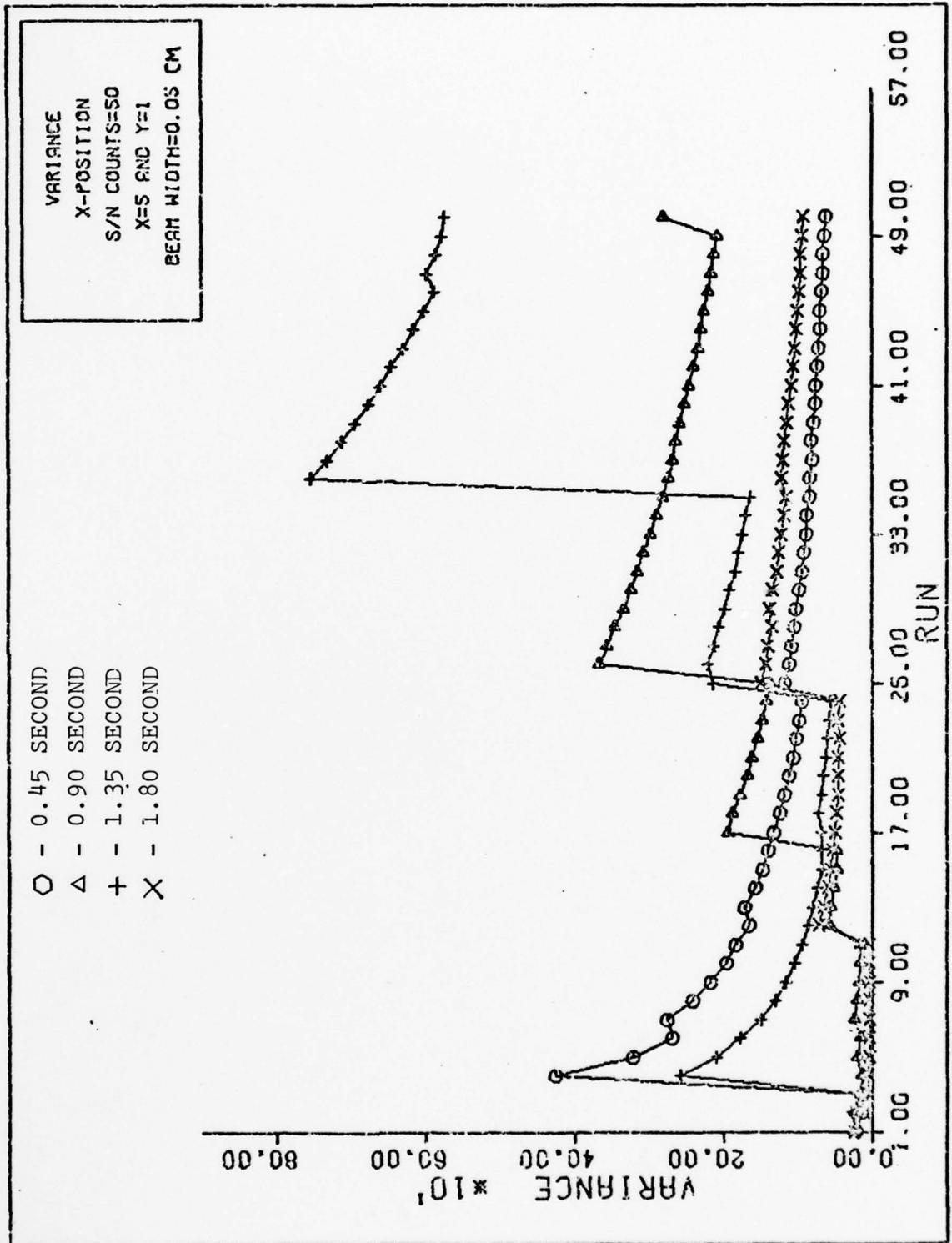Figure 24d X-POSITION VARIANCES

111

Figure 25a MEAN CONVERGENCE

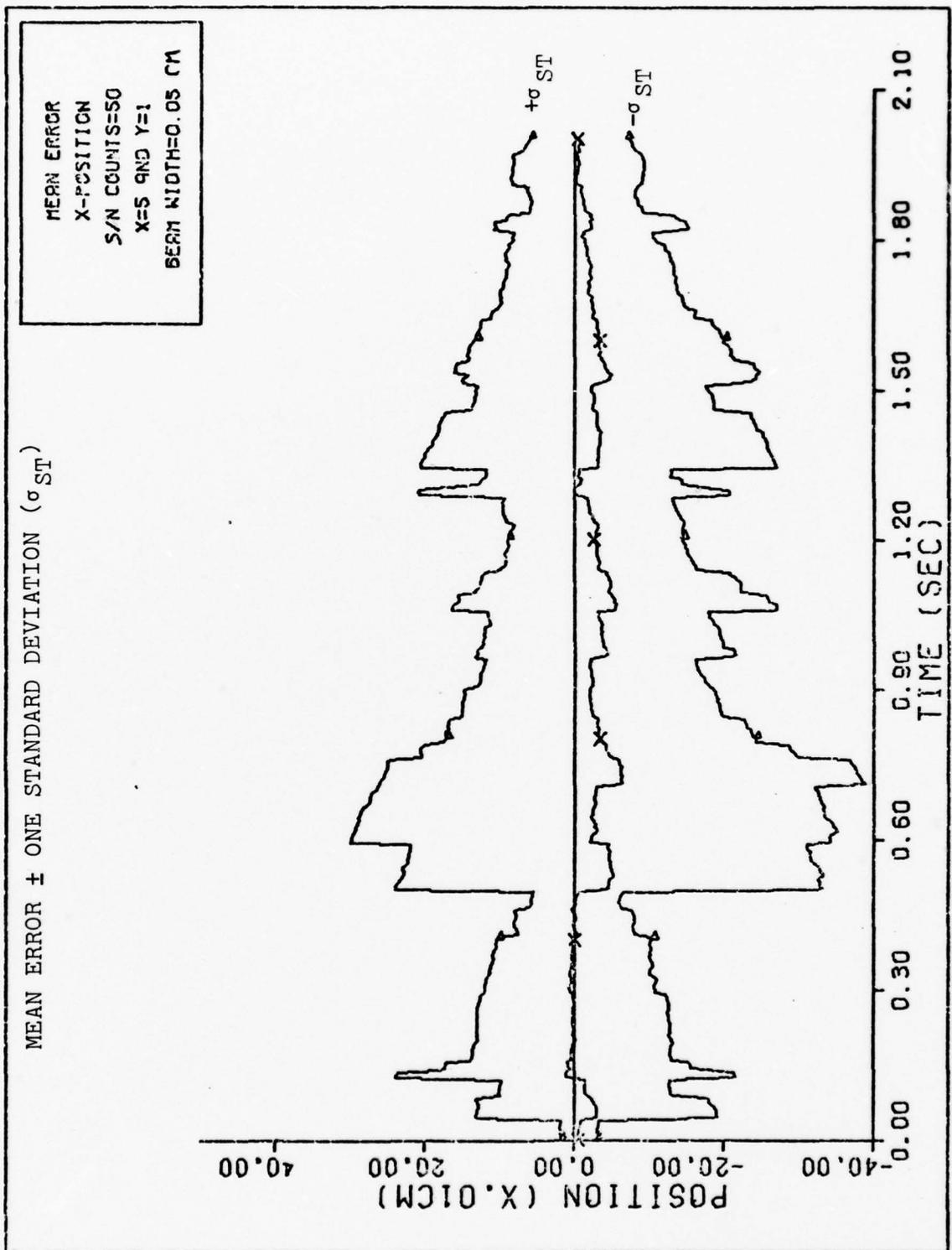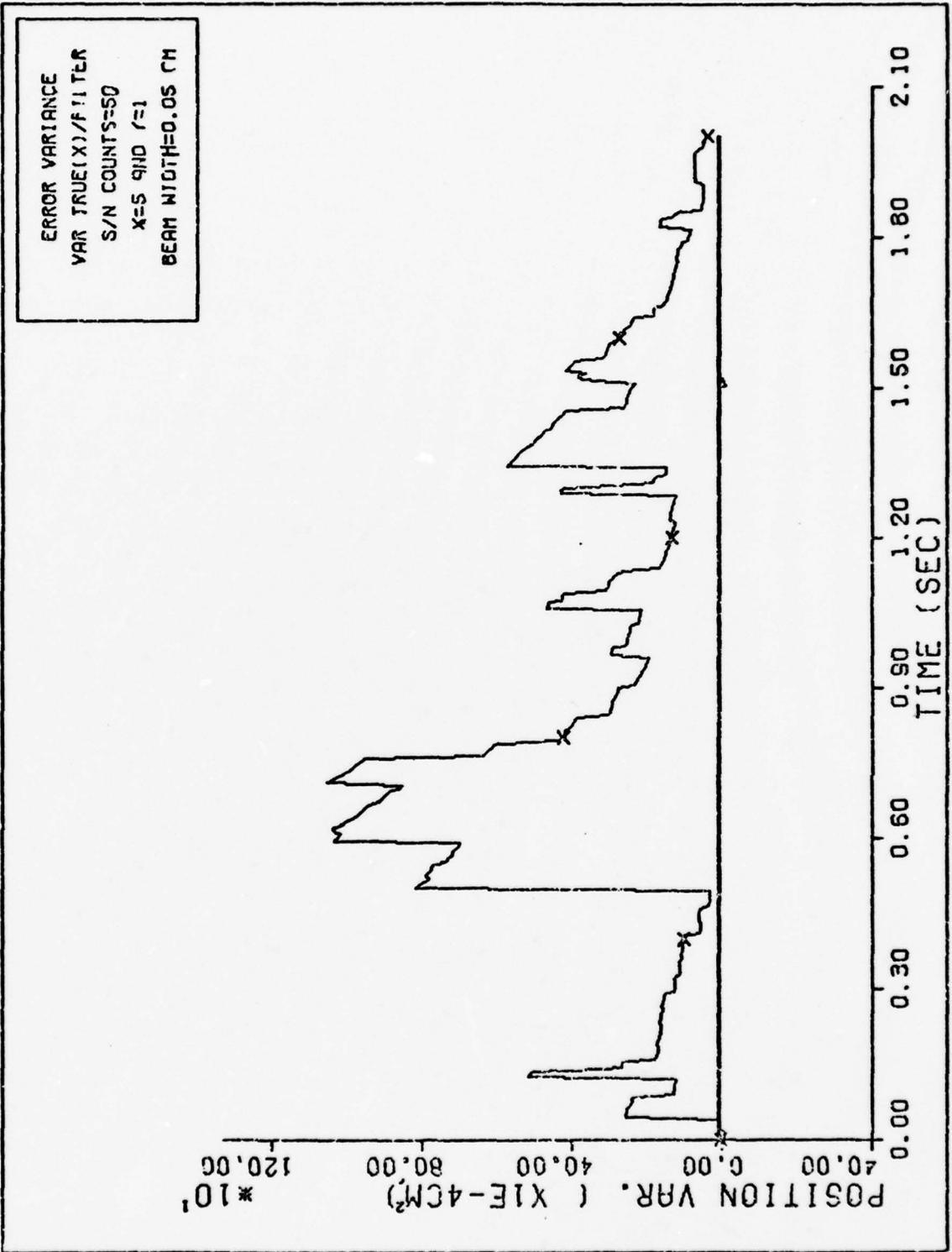Figure 25b   VARIANCE CONVERGENCE

113

Figure 25c X-POSITION
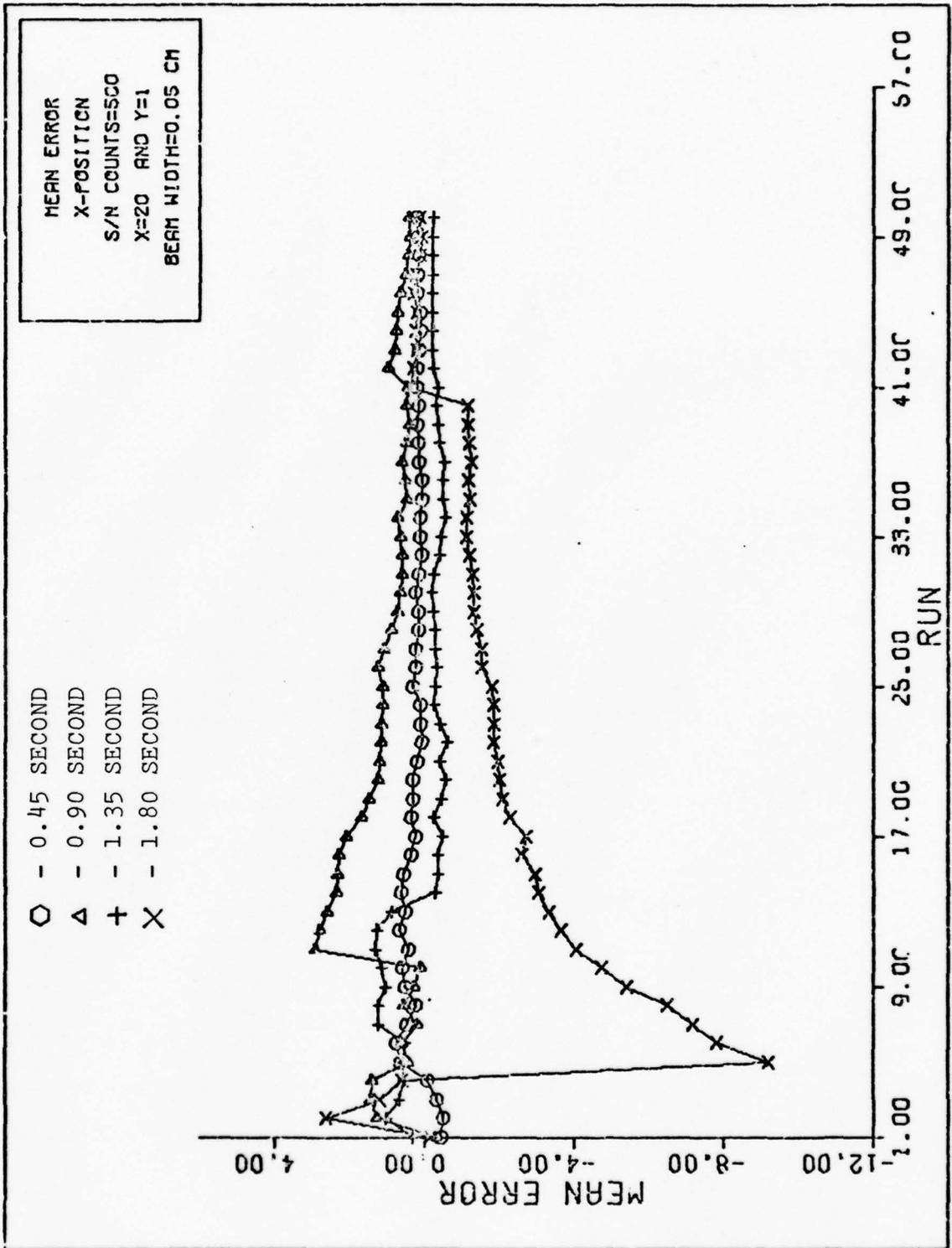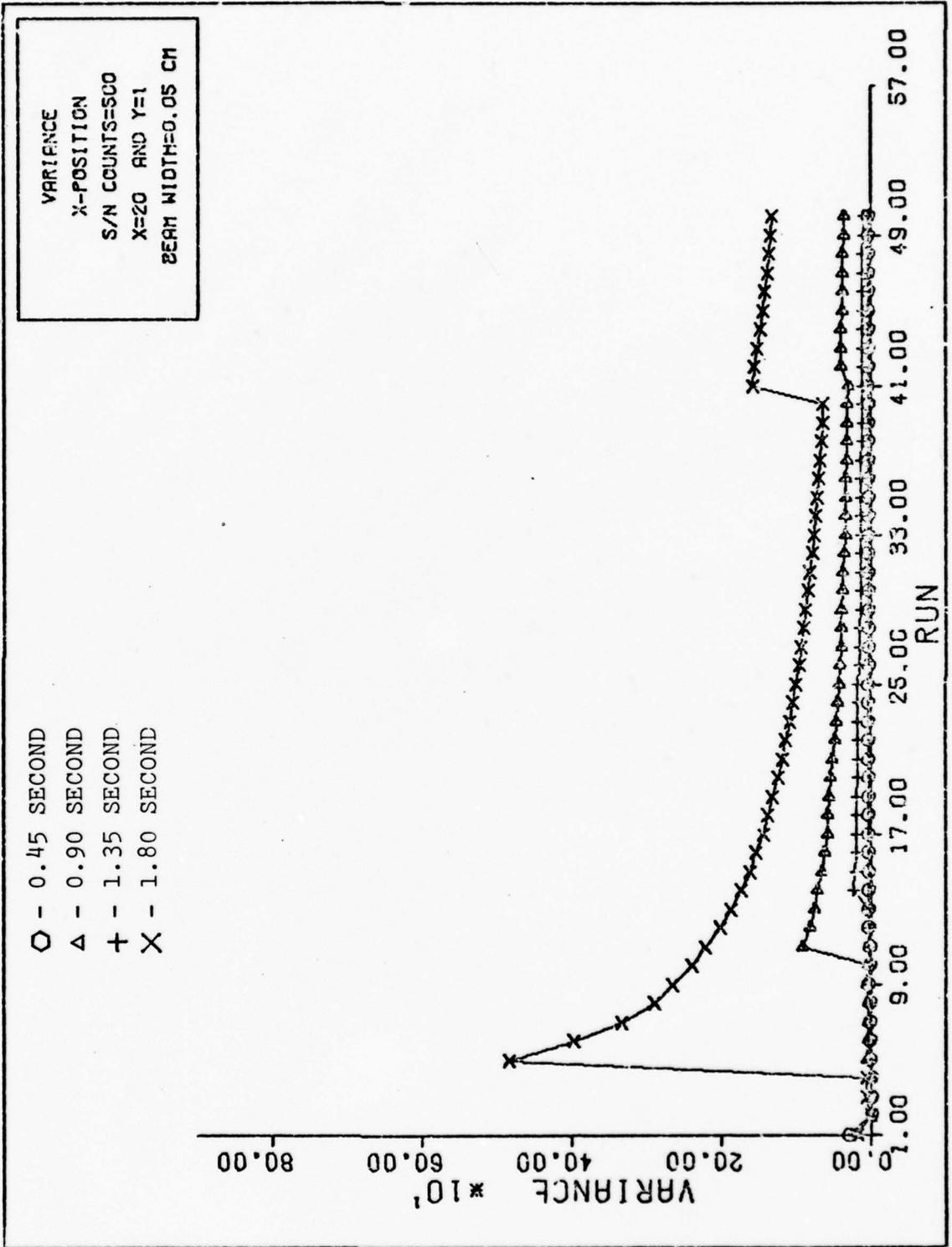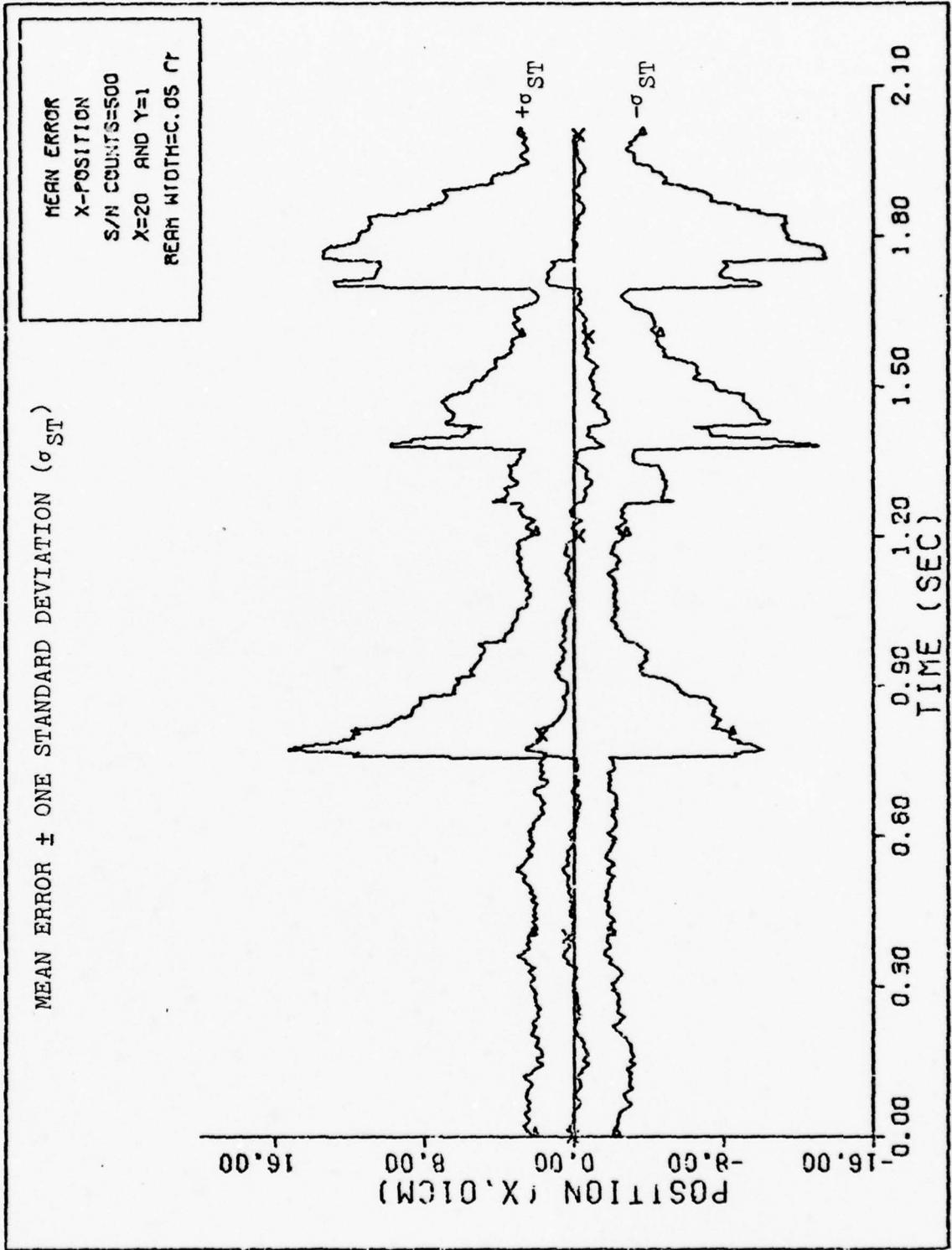
114

Figure 25d X-POSITION VARIANCES

115

Figure 26a MEAN CONVERGENCE

116
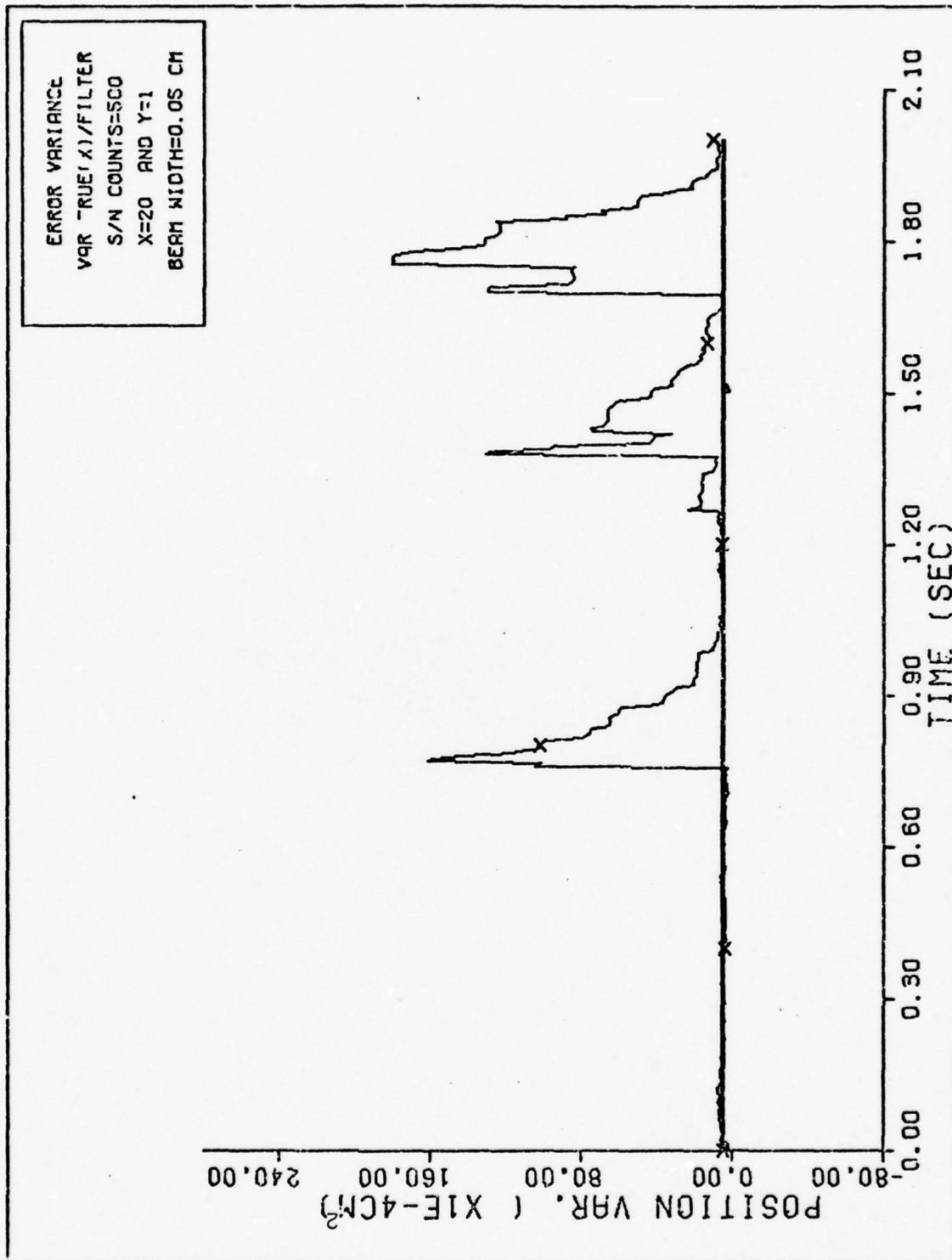
Figure 26b VARIANCE CONVERGENCE

Figure 26c X-POSITION

118

Figure 26d X-POSITION VARIANCES

Figure 27a MEAN CONVERGENCE

120

Figure 27b VARIANCE CONVERGENCE

121

Figure 27c X-POSITION

122

Figure 27d X-POSITION VARIANCES

123

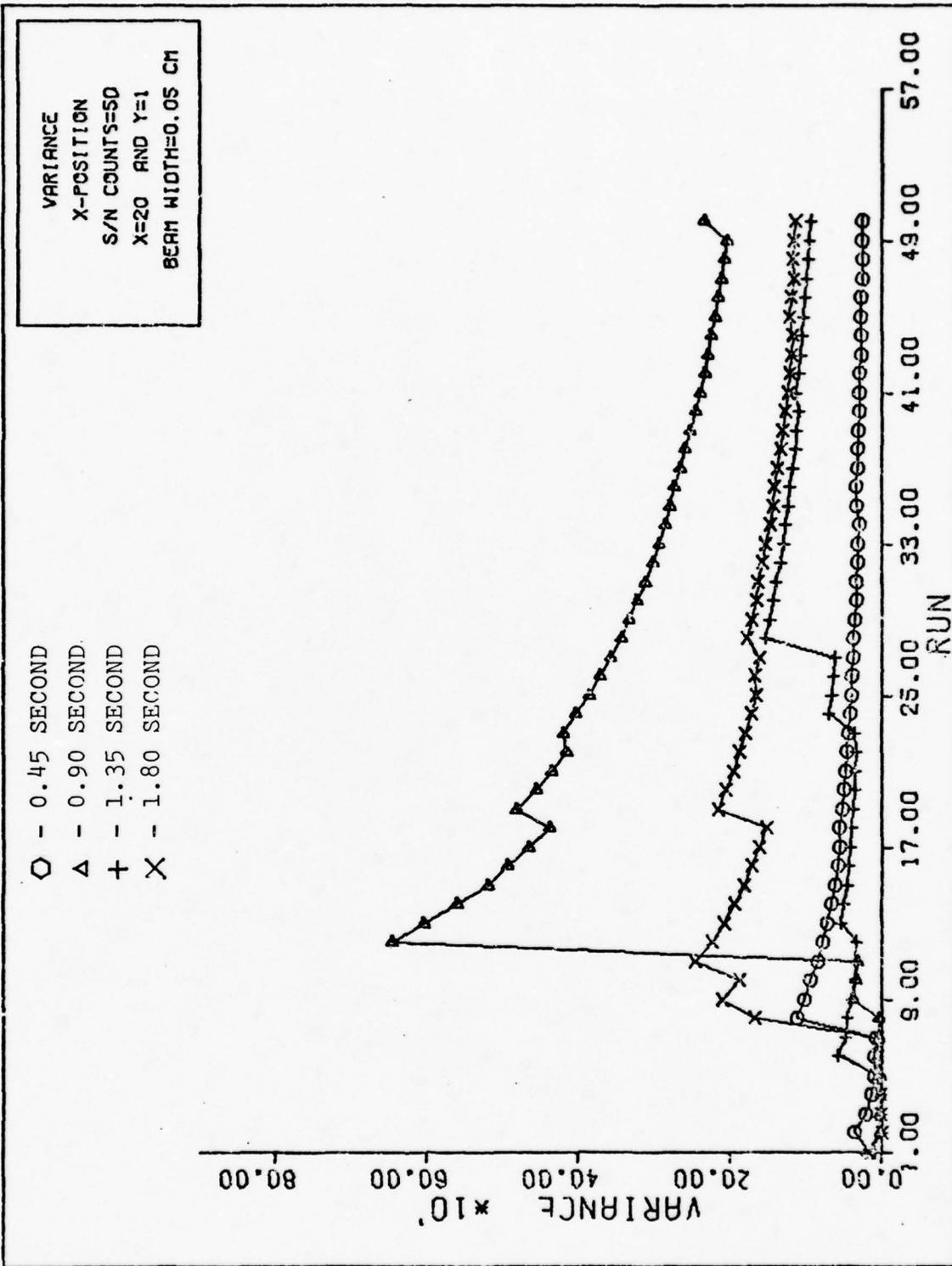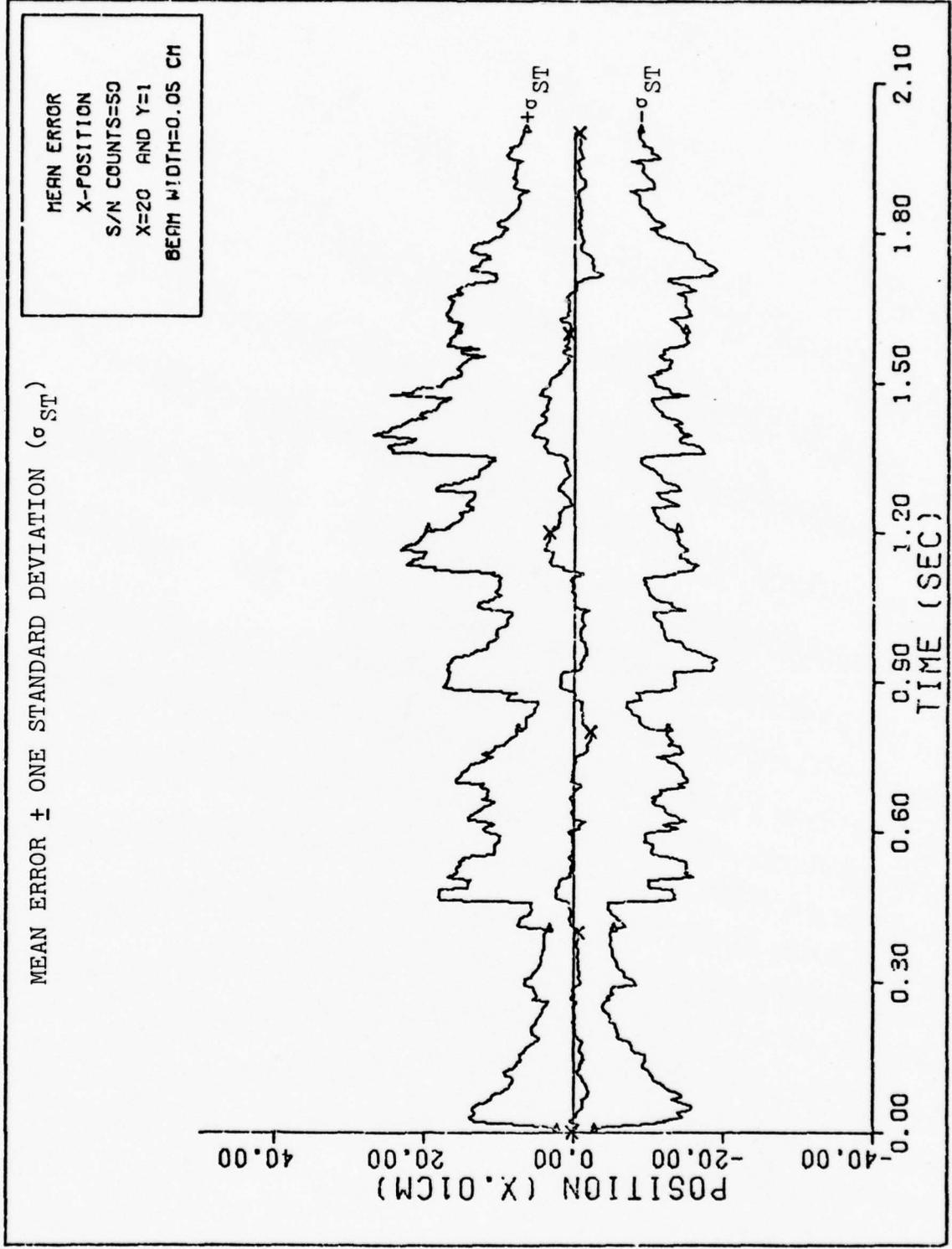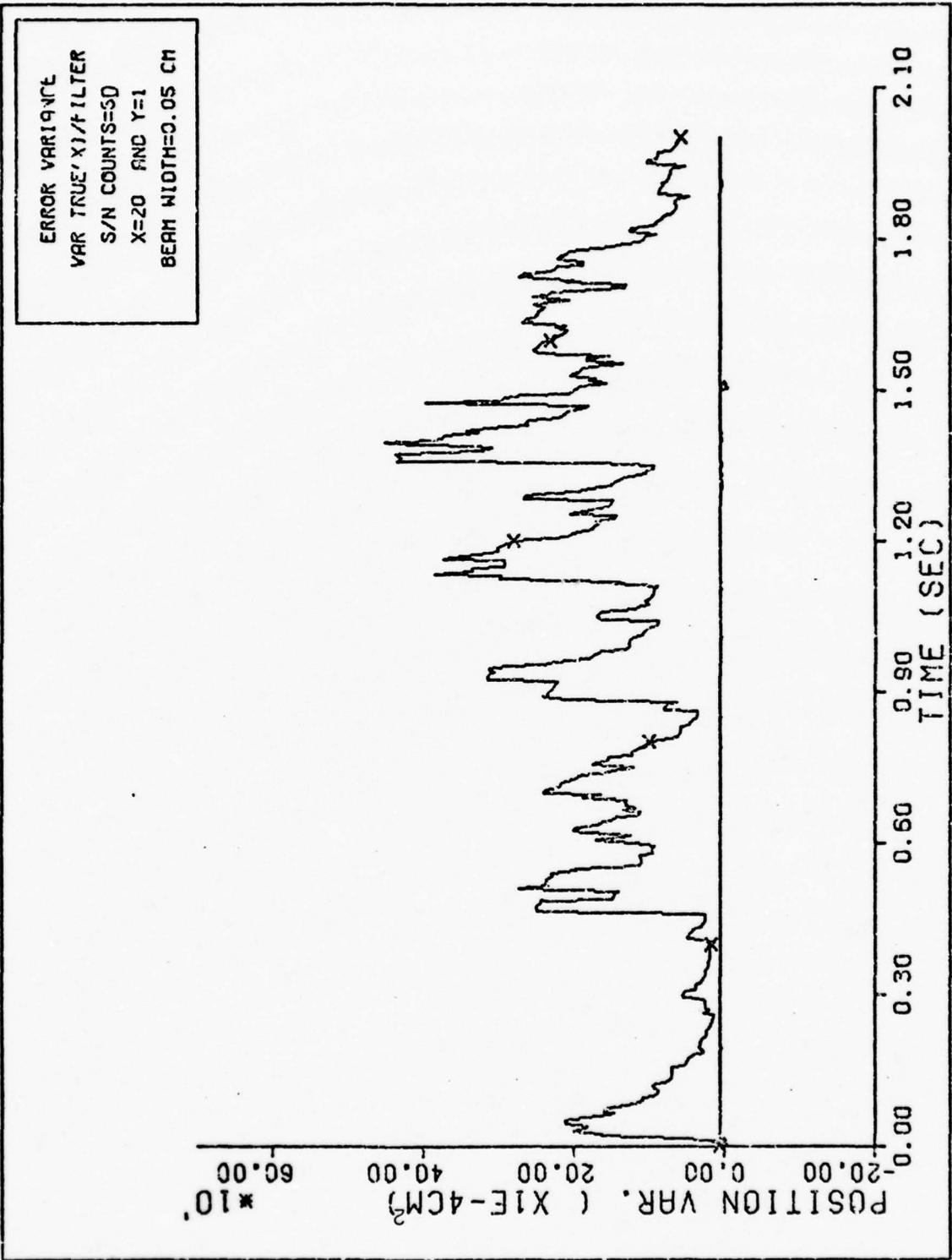Figure 28a MEAN CONVERGENCE

124

Figure 28b VARIANCE CONVERGENCE

Figure 28c X-POSITION

126

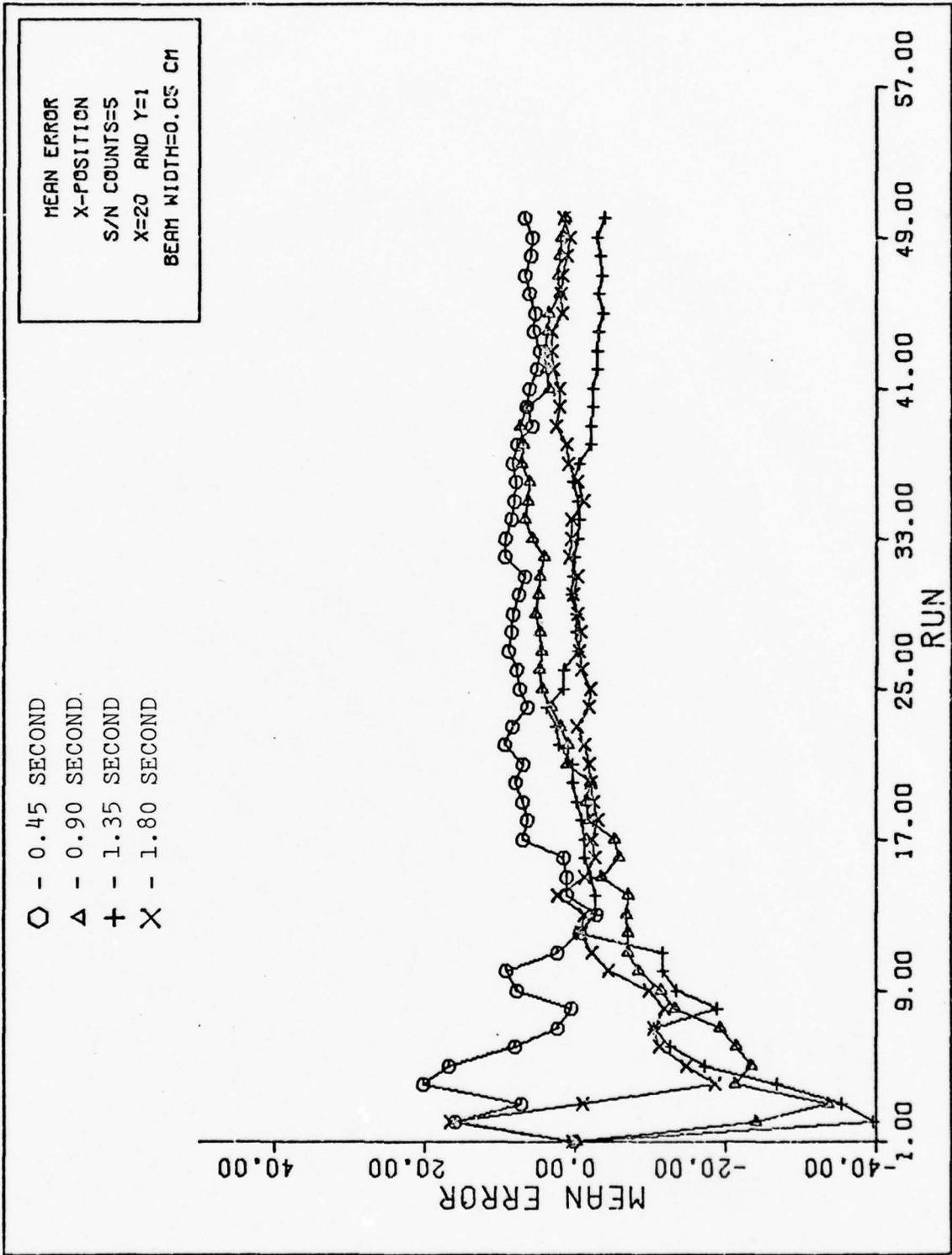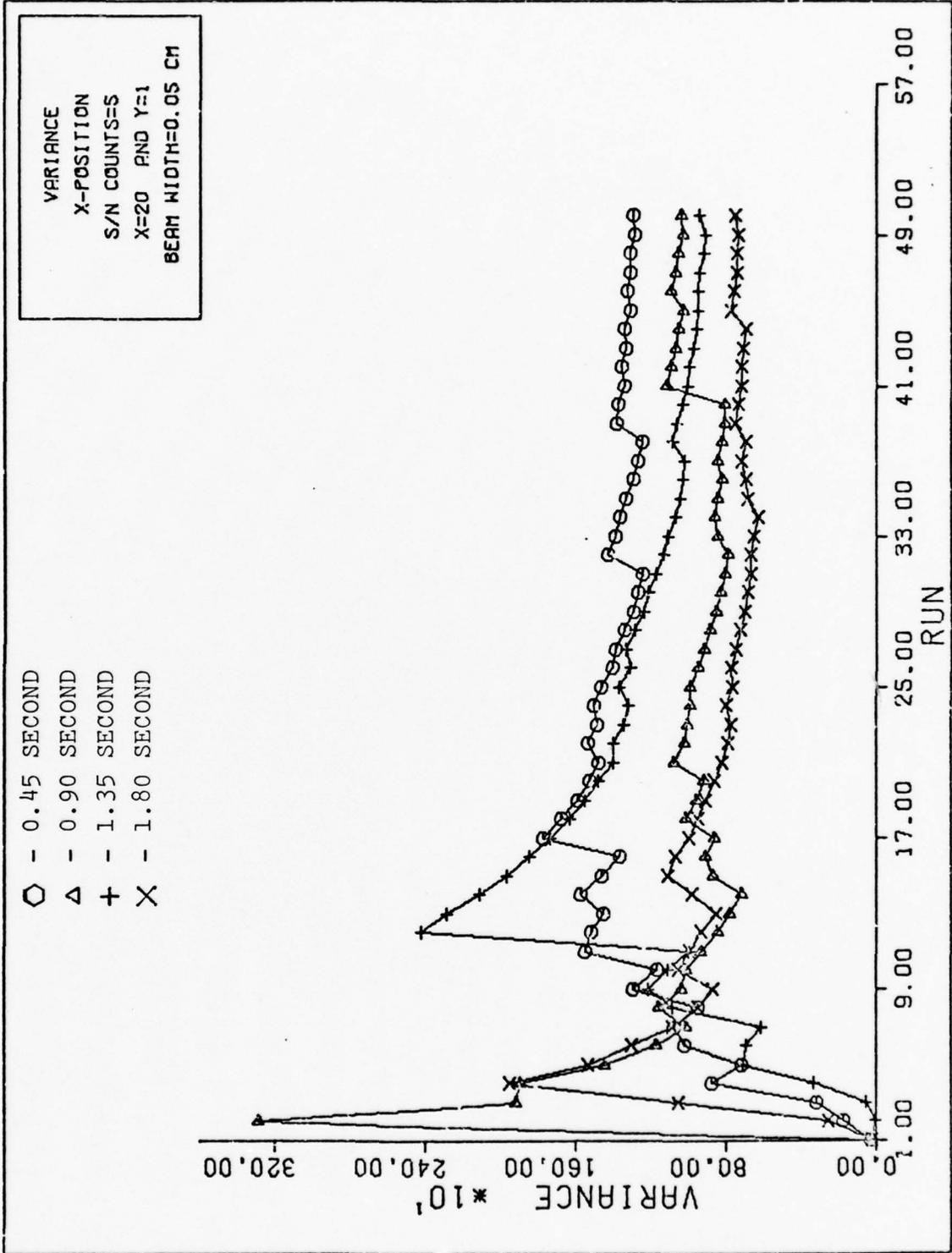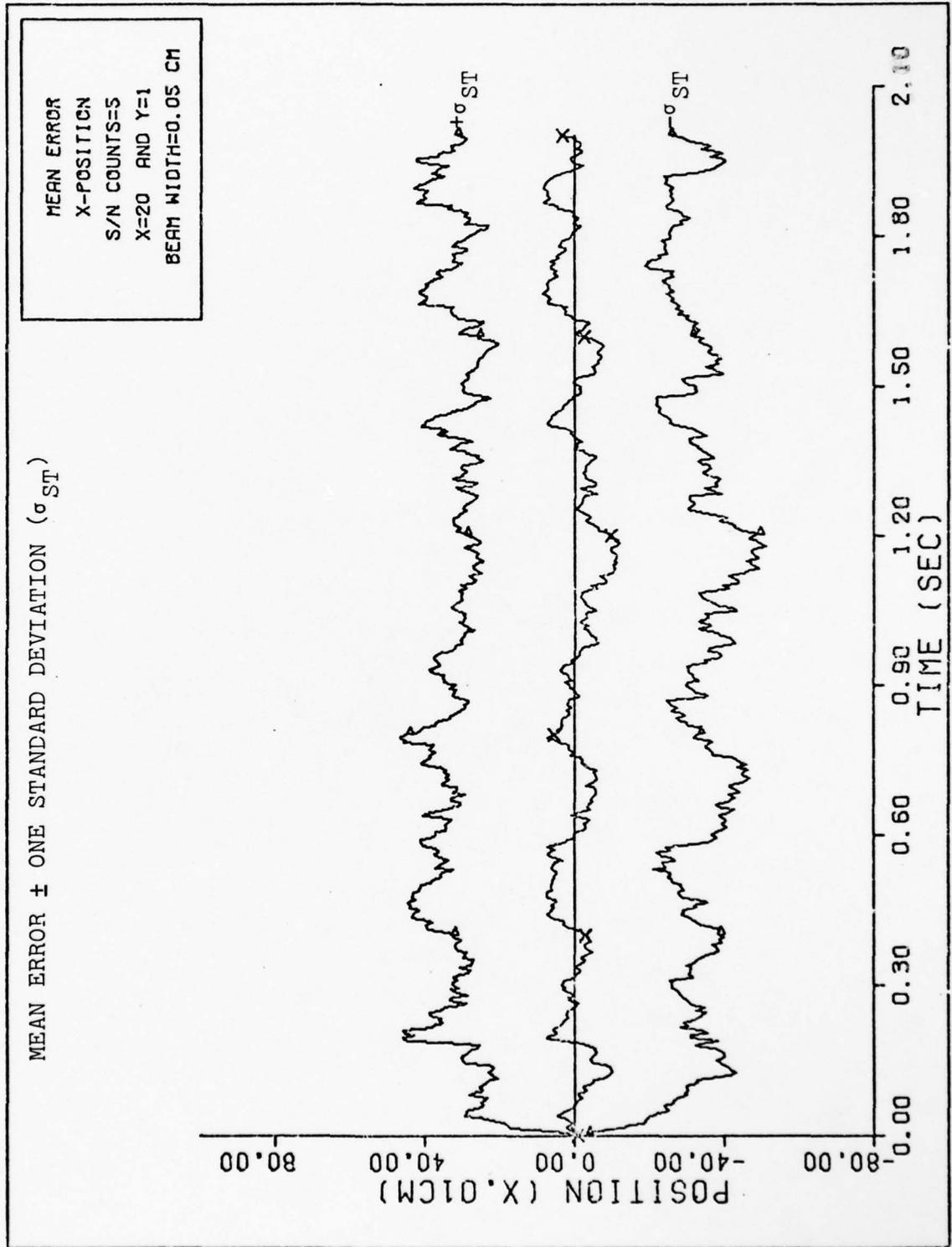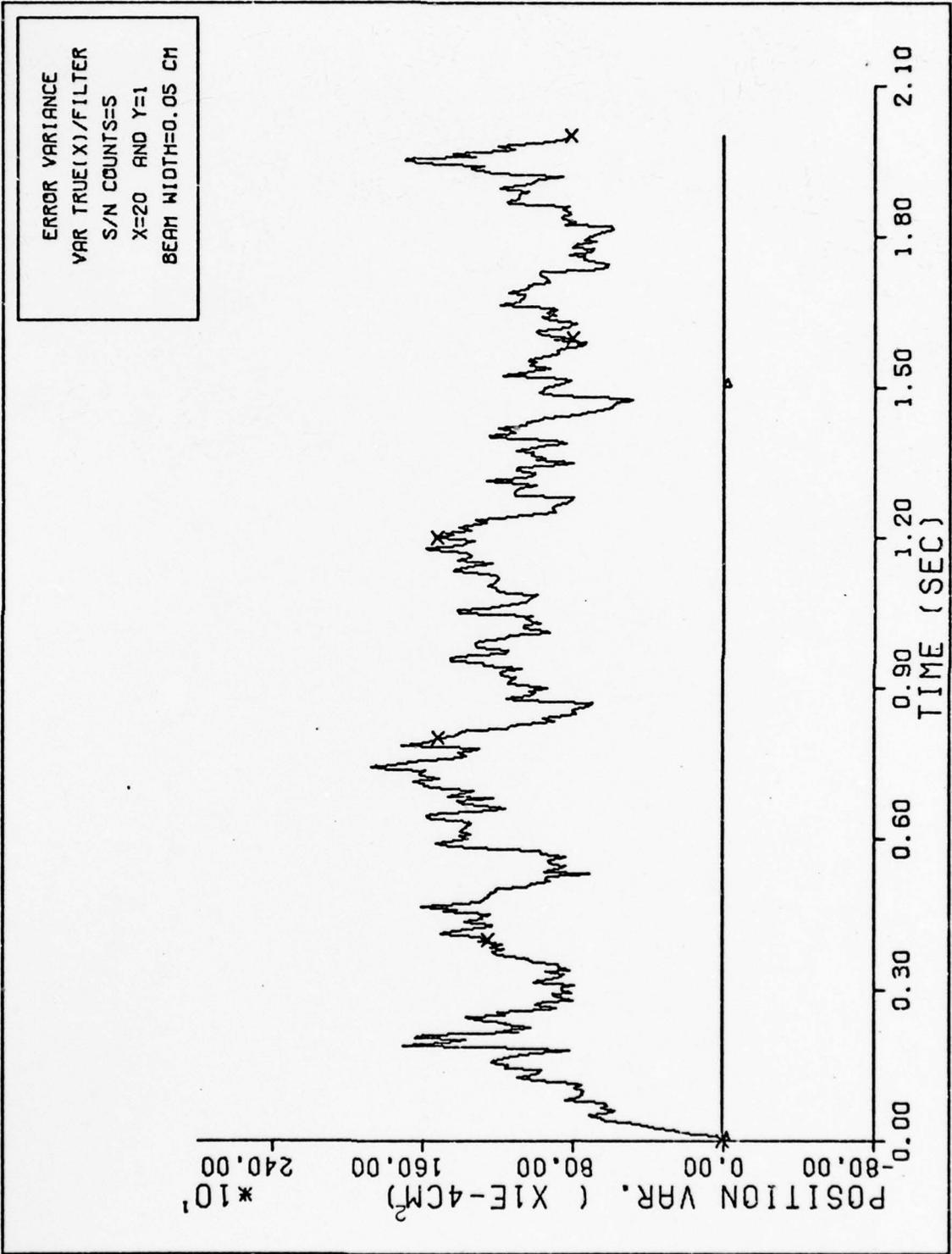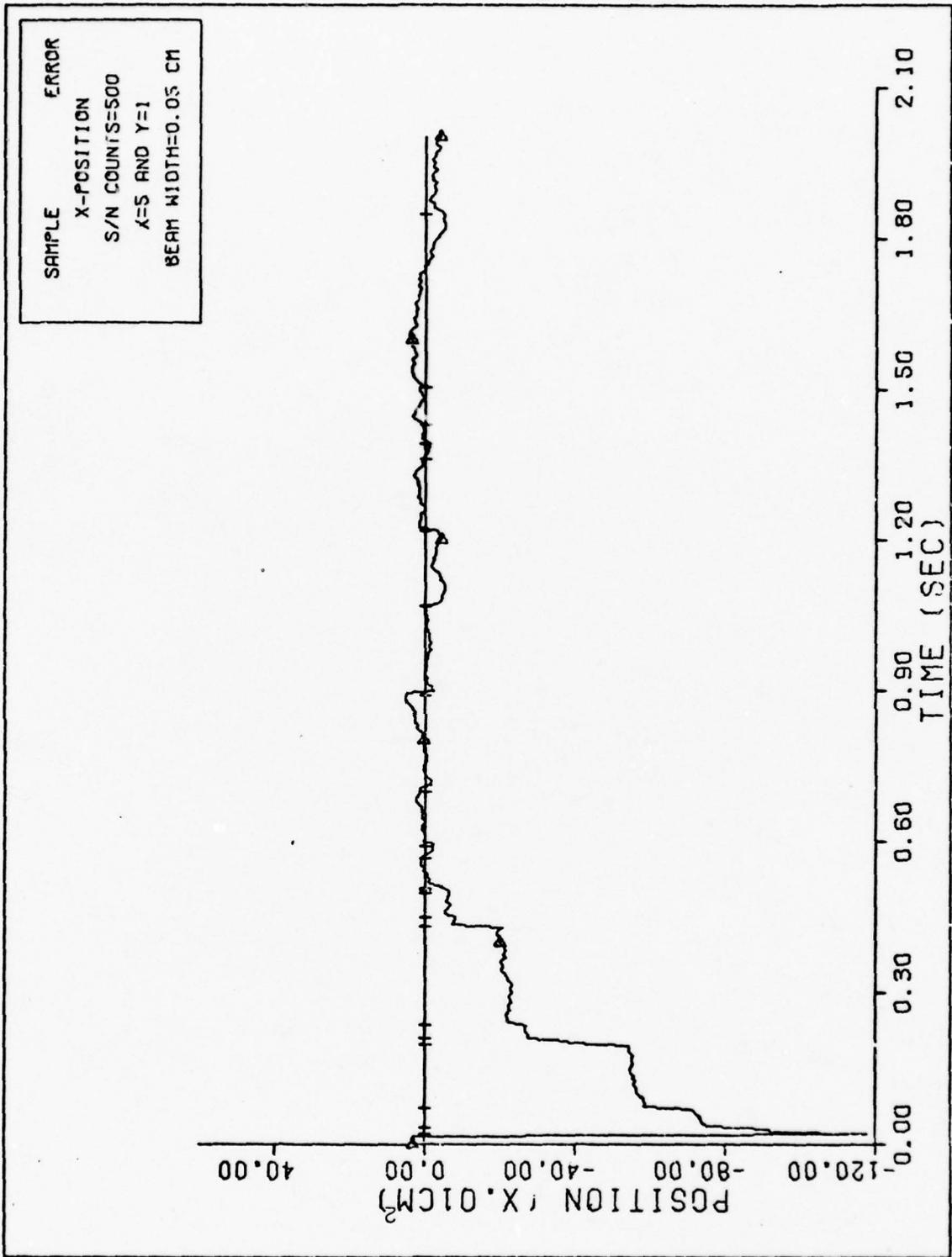Figure 28d X-POSITION VARIANCES

127

Figure 29a X-POSITION

128

Figure 29b X-POSITION FILTER VAR'ANCE

129
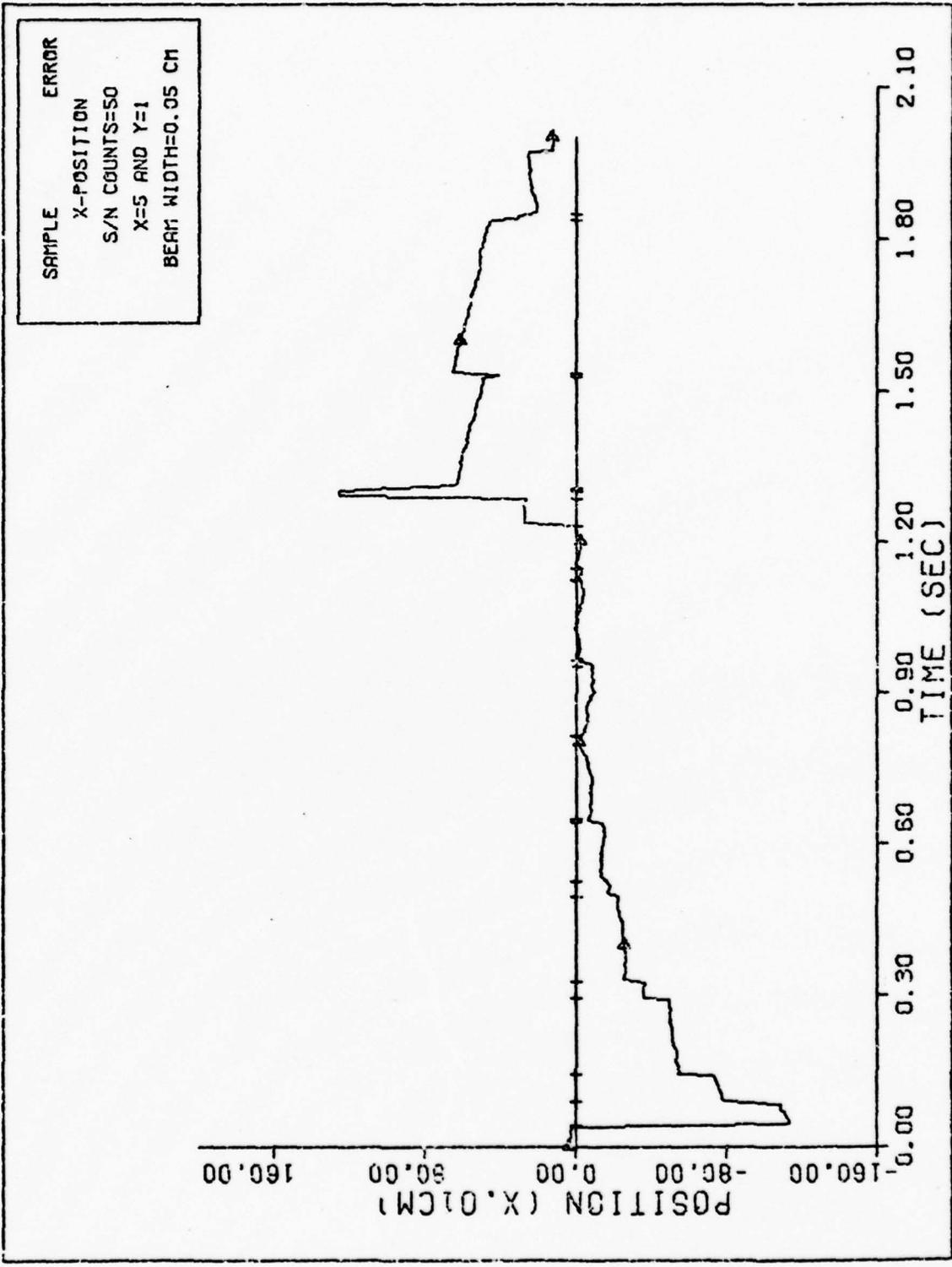
Figure 30a   X-POSITION

130

Figure 30b X-POSITION FILTER VARIANCE
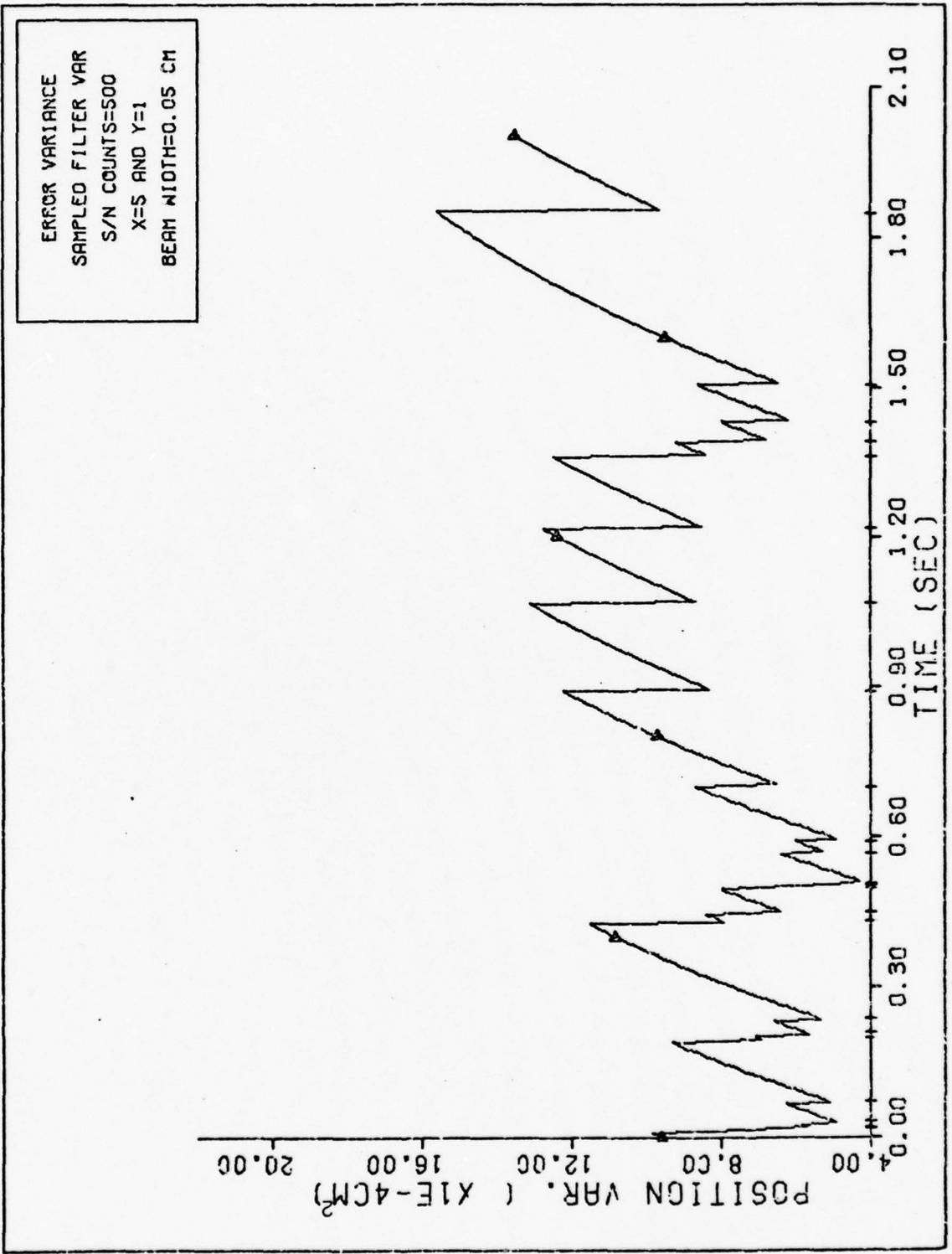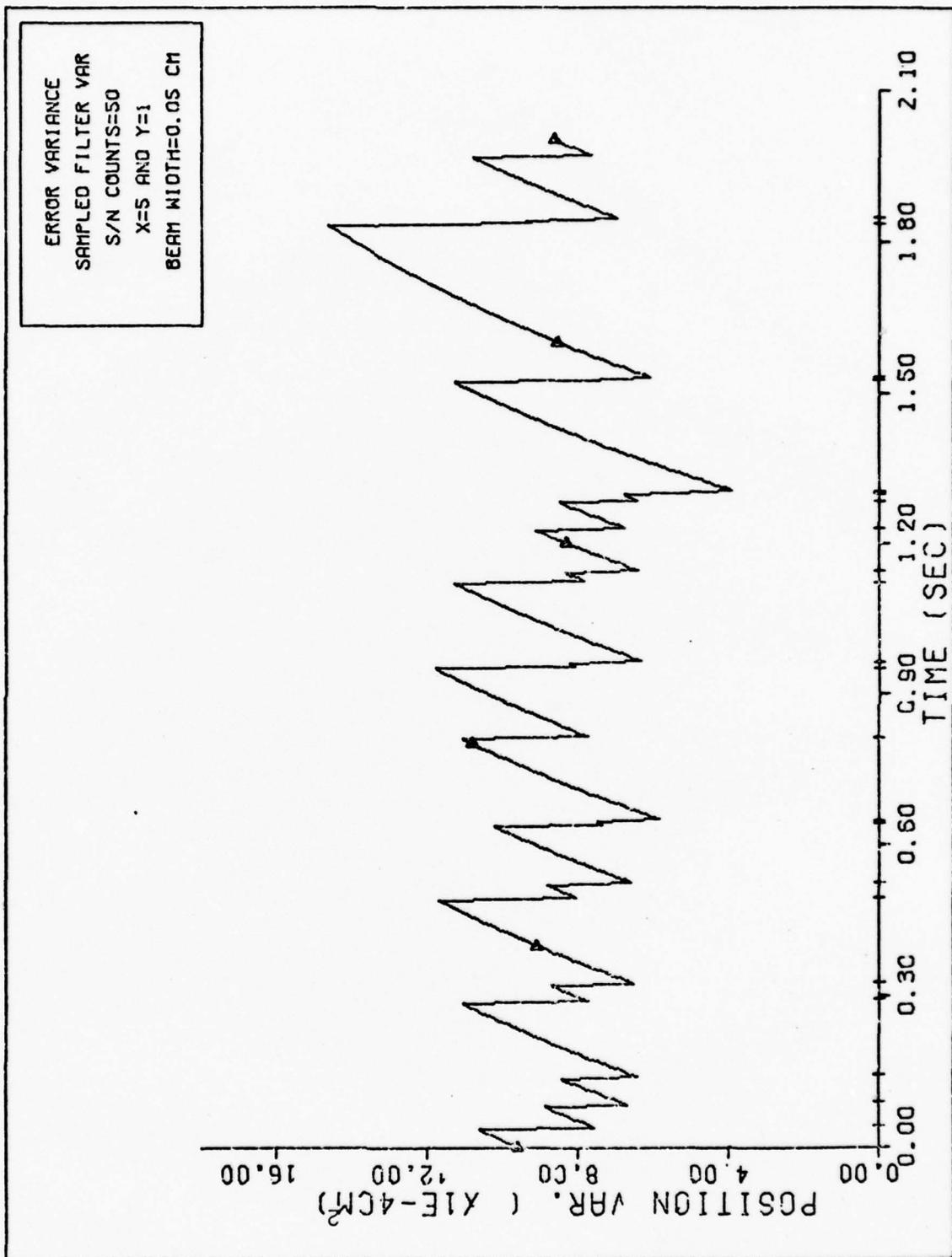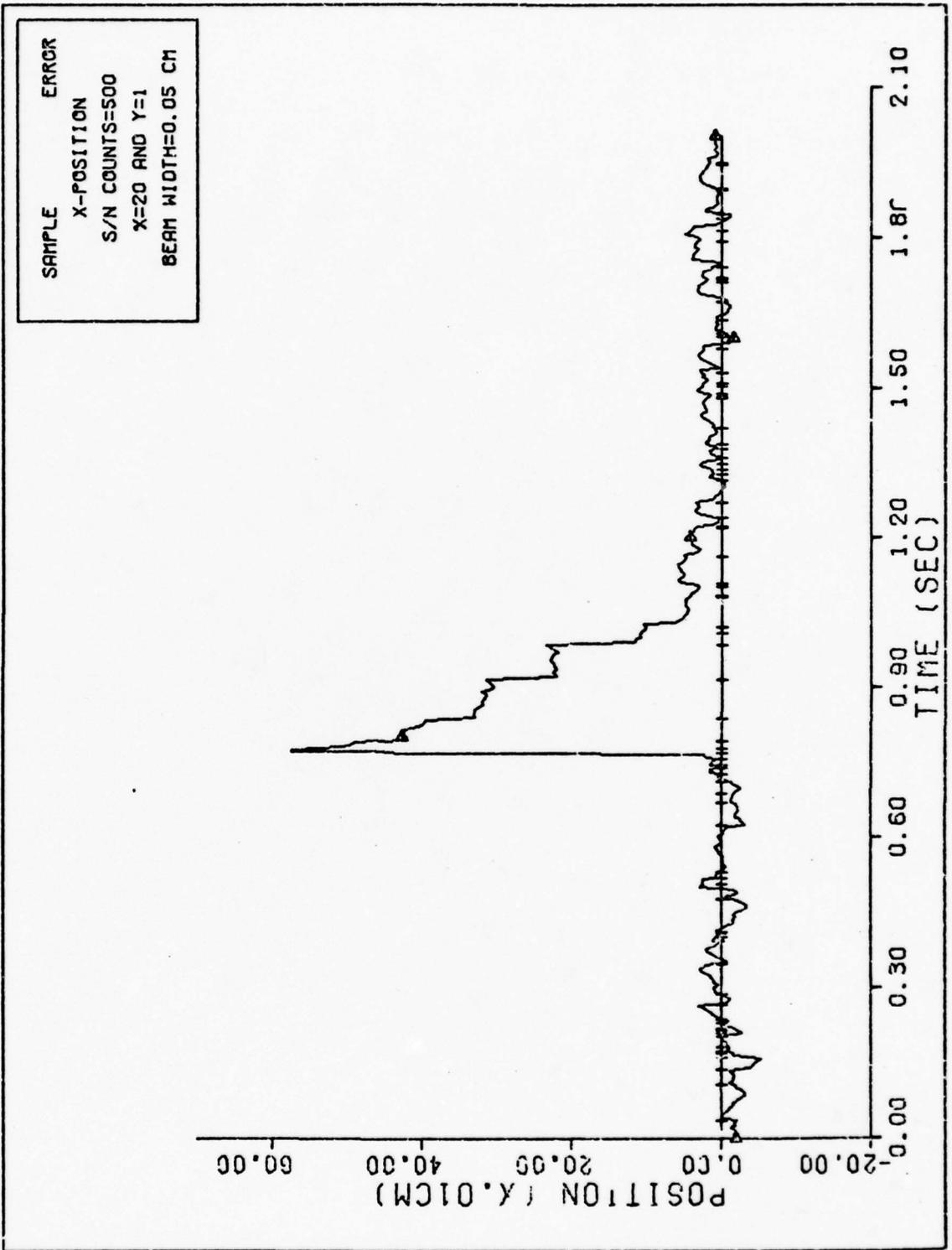
131

Figure 31a X-POSITION

132
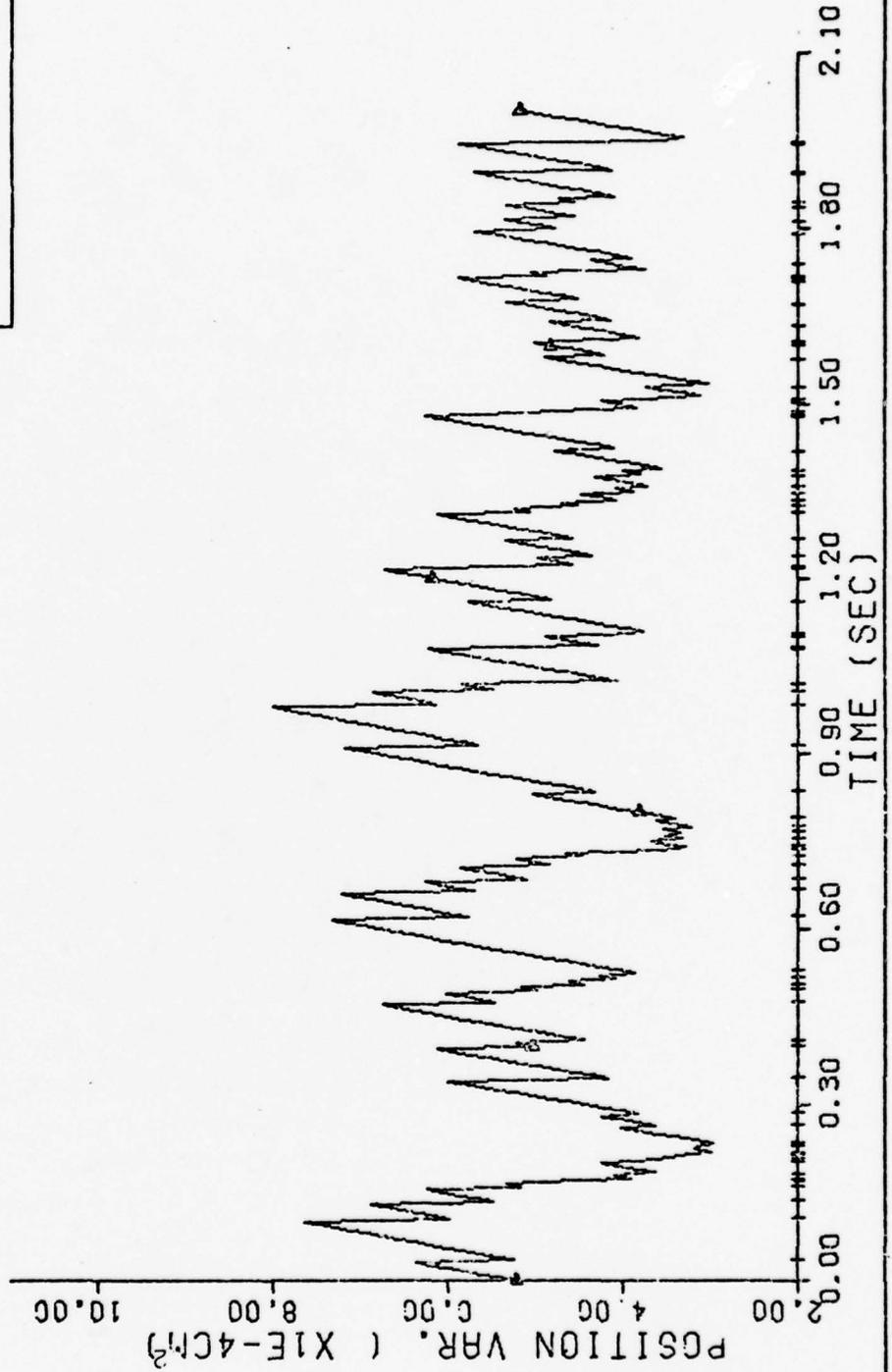
Figure 31b X-POSITION FILTER VARIANCE

133
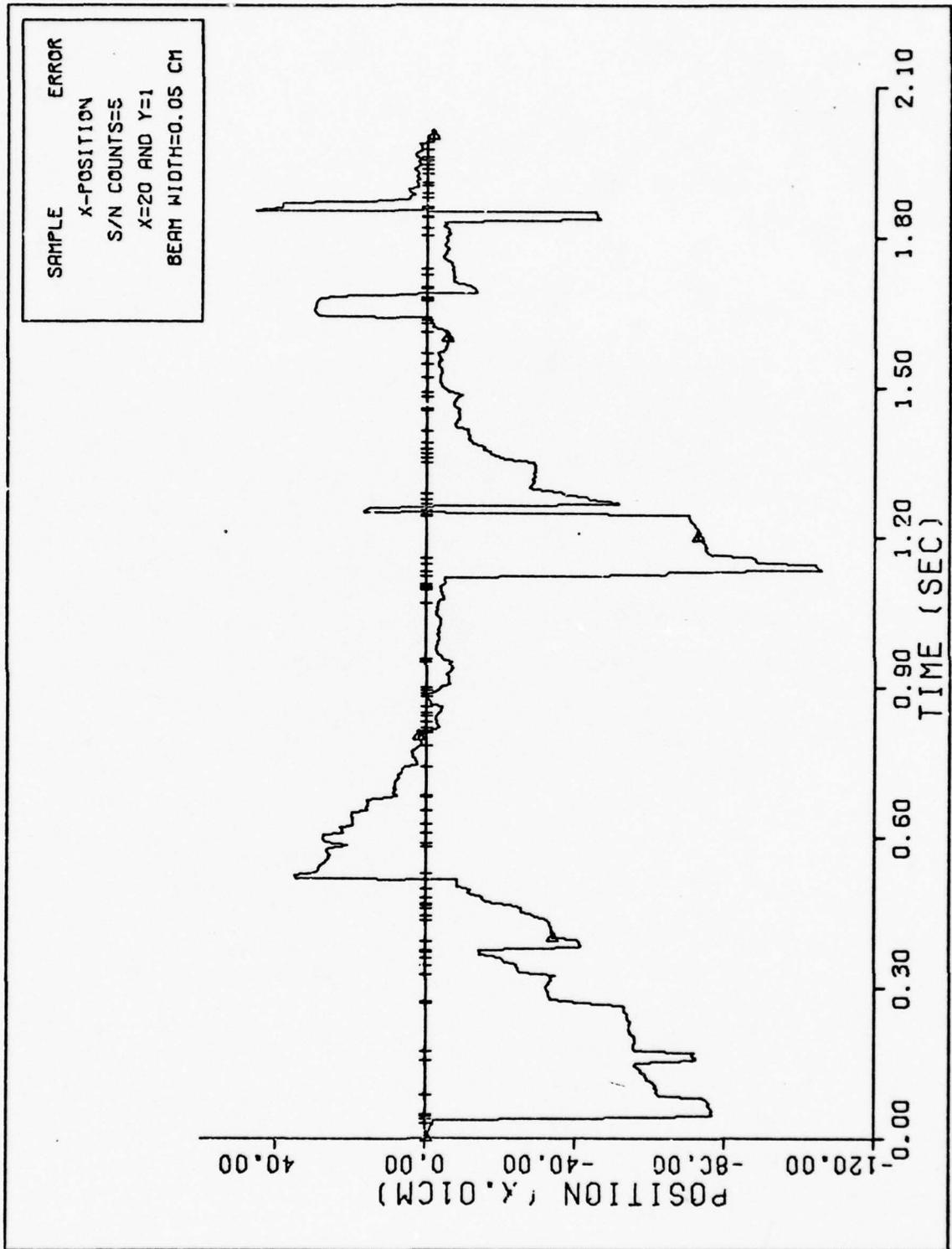
Figure 32a X-POSITION
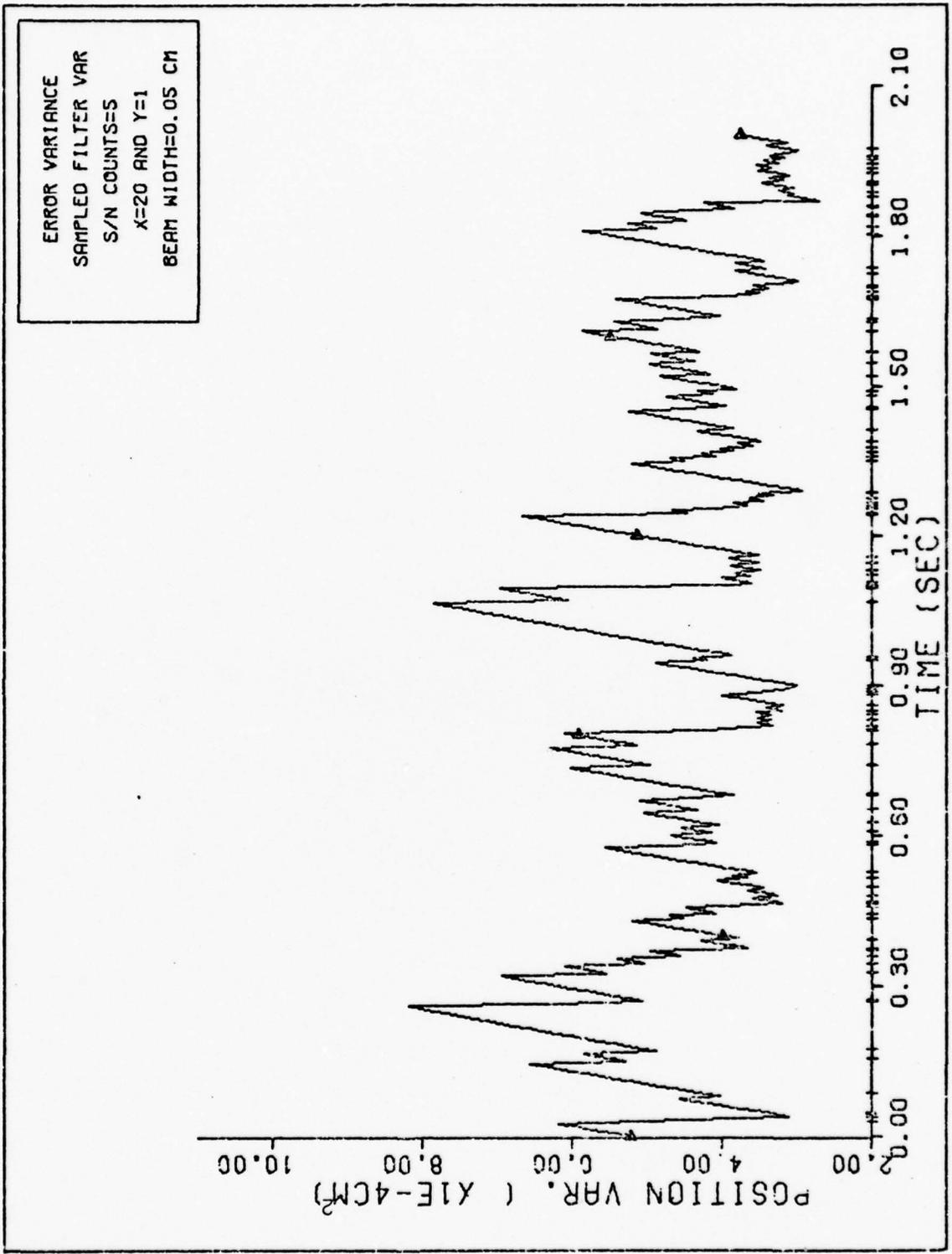
Figure 32b X-POSITION FILTER VARIANCE

135

the beam width is decreased, which means that the signal events
are distributed at locations very near the position of maximum
intensity, then the filter should weigh the measurement heavily
since there is less "corruption" in the measurements. Recall
that the filter assumes the photodetector array is very large;
and for dark limited conditions, the dark events can occur any-
where in the array with equal probability. Thus, the beam width
must be very wide in order to put less weight in the dark measure-
ments. Since the filter cannot distinguish between a dark or
signal event, the signal measurements will also have less weight.
In other words, the beam is assumed to be wide in order to "wash"
out the dark events; and the signal events are also assumed to
be distributed at larger distances from the spot's centroid of
maximum intensity. From this discussion, it becomes evident that
altering the beam width in order to compensate for the dark
events has some problems, especially if the photodector array is
dark current-limited. It would be preferable to try to distinguish
signal events from dark events and weigh them properly.

In residual monitoring, the sequence of residuals are
monitored in real-time to decide whether the current measurement
value can be discarded as a "bad" data point. Consider Eq. (23)
which is repeated below as

$$\hat{x}(t_k+) = \hat{x}(t_k-) + K(t_k-)[\underbrace{r_k-\hat{x}(t_k-)}_{\text{residual}}] \tag{65}$$

where $H(t_k)=1$. From Eq. (65), a residual is simply the
difference between the measurement, r_k, and the conditional

mean $\hat{x}(t_k-)$ of the measurement, conditioned on only the previous history of measurements. Residual monitoring would be used to test for events that are not due to the spot. If the events fail the test, then the filter does not process the dark measurement. Instead, the filter propagates its conditional mean and covariance until it senses another event which passes the test. For this study, if the residual in Eq. (23) is greater than three times the beam width, then the event is discarded. Note that in residual monitoring, there is no need to increase the beam width. In other words, the event locations (due to the spot) have the same distribution about the spot's centroid, as before. Thus, residual monitoring can be exploited effectively for high signal-to-noise counts ratio, since few events will fail the test. Physically, residual monitoring is a better method of compensation than filter tuning in order to handle the dark current effects.

As a rough comparison in performance improvement between the above methods, the square of the mean errors at the sample times are all summed. The results from this method are listed in Tables 3 and 4. The tables indicate that residual monitoring is a better method to handle the dark current effects than altering the beam width. For particular cases, plots of the variances in which these methods are employed are shown in Figs. (33) through (36). These plots show that the Snyder Filter performance can be enhanced substantially through residual monitoring by not processing the dark measurement. As expected, these results correspond to the physical arguments mentioned in previous paragraphs.

137

Average Signal-to-Noise Counts Ratio	Method of Compensation	Beam Width (x0.01 cm)	Sum of the Square of the Mean Error
∞	None	5.0	207.0
500	None	5.0	555.0
500	Residual Monitoring	5.0	80.2
500	Filter Tuning	7.5	323.0
500	Filter Tuning	10.0	223.0
500	Filter Tuning	15.0	166.0
500	Filter Tuning	20.0	167.0
500	Filter Tuning	30.0	196.0
50	None	5.0	3150.0
50	Residual Monitoring	5.0	66.3
50	Filter Tuning	10.0	1230.0
50	Filter Tuning	20.0	198.0
50	Filter Tuning	40.0	110.0
50	Filter Tuning	20.0	377.0
1	None	5.0	2860.0
1	Residual Monitoring	5.0	33.0
1	Filter Tuning	30.0	350.0
1	Filter Tuning	40.0	70.7

Table 4. Test Cases for X=5 and Y=1
(with dark current)

138

Average Signal-to-Noise Counts Ratio	Method of Compensation	Beam Width (x0.01 cm)	Sum of the Square of the Mean Error
∞	None	5.0	42.7
500	None	5.0	171.3
500	Residual Monitoring	5.0	33.6
500	Filter Tuning	6.0	142.4
500	Filter Tuning	7.0	122.3
500	Filter Tuning	8.0	95.1
500	Filter Tuning	10.0	83.7
500	Filter Tuning	20.0	57.8
500	Filter Tuning	30.0	50.9
50	None	5.0	1160.0
50	Residual Monitoring	5.0	41.0
50	Filter Tuning	10.0	423.0
50	Filter Tuning	15.0	208.0
50	Filter Tuning	30.0	97.2
50	Filter Tuning	40.0	105.0
5	None	5.0	8940.0
5	Residual Monitoring	5.0	60.6
5	Filter Tuning	20.0	882.0
5	Filter Tuning	30.0	396.0

Table 5. Test Cases for X=20 and Y=1
(with dark current)
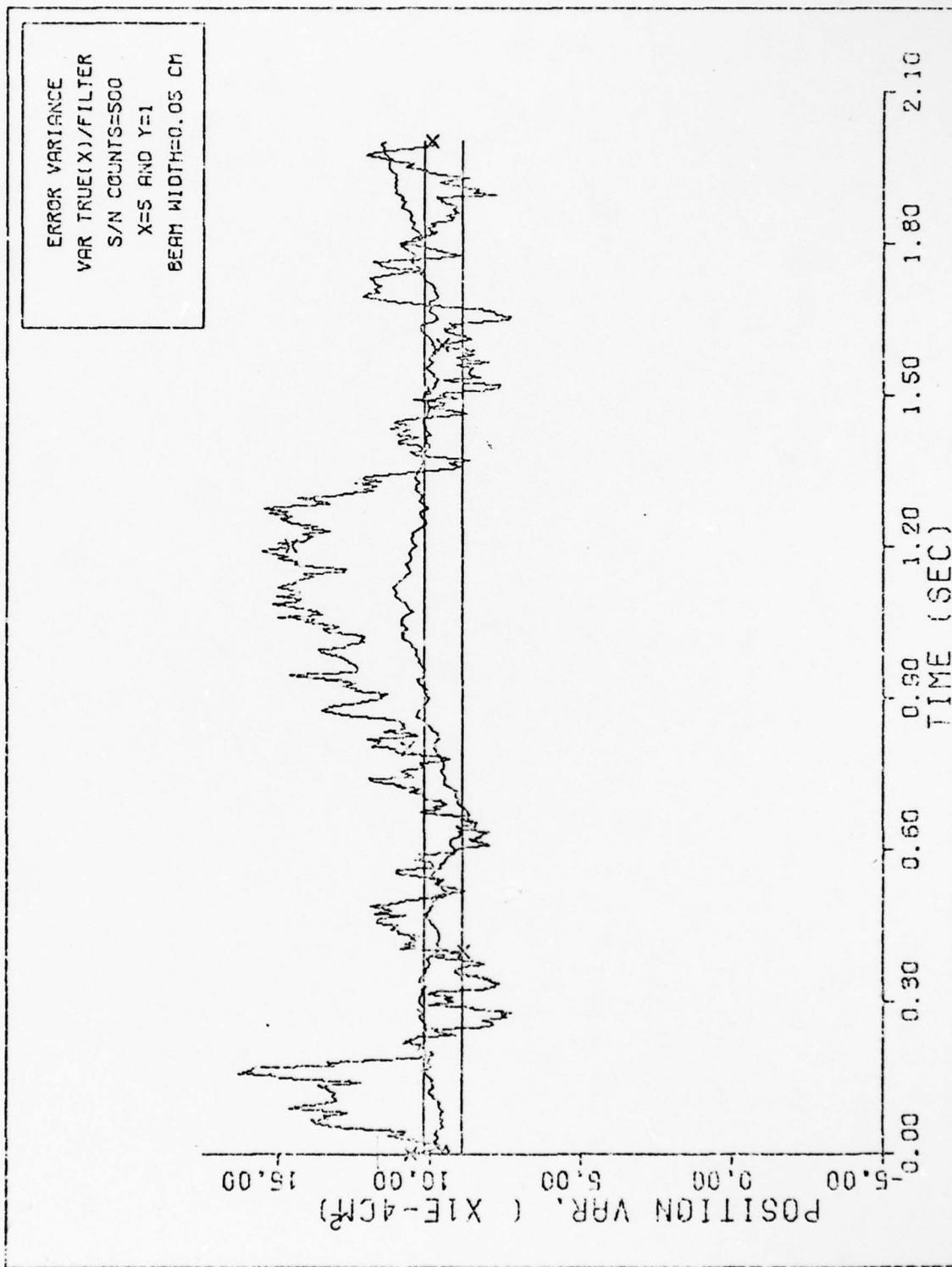
Figure 33a X-POSITION VARIANCES
(RESIDUAL MONITORING)

140
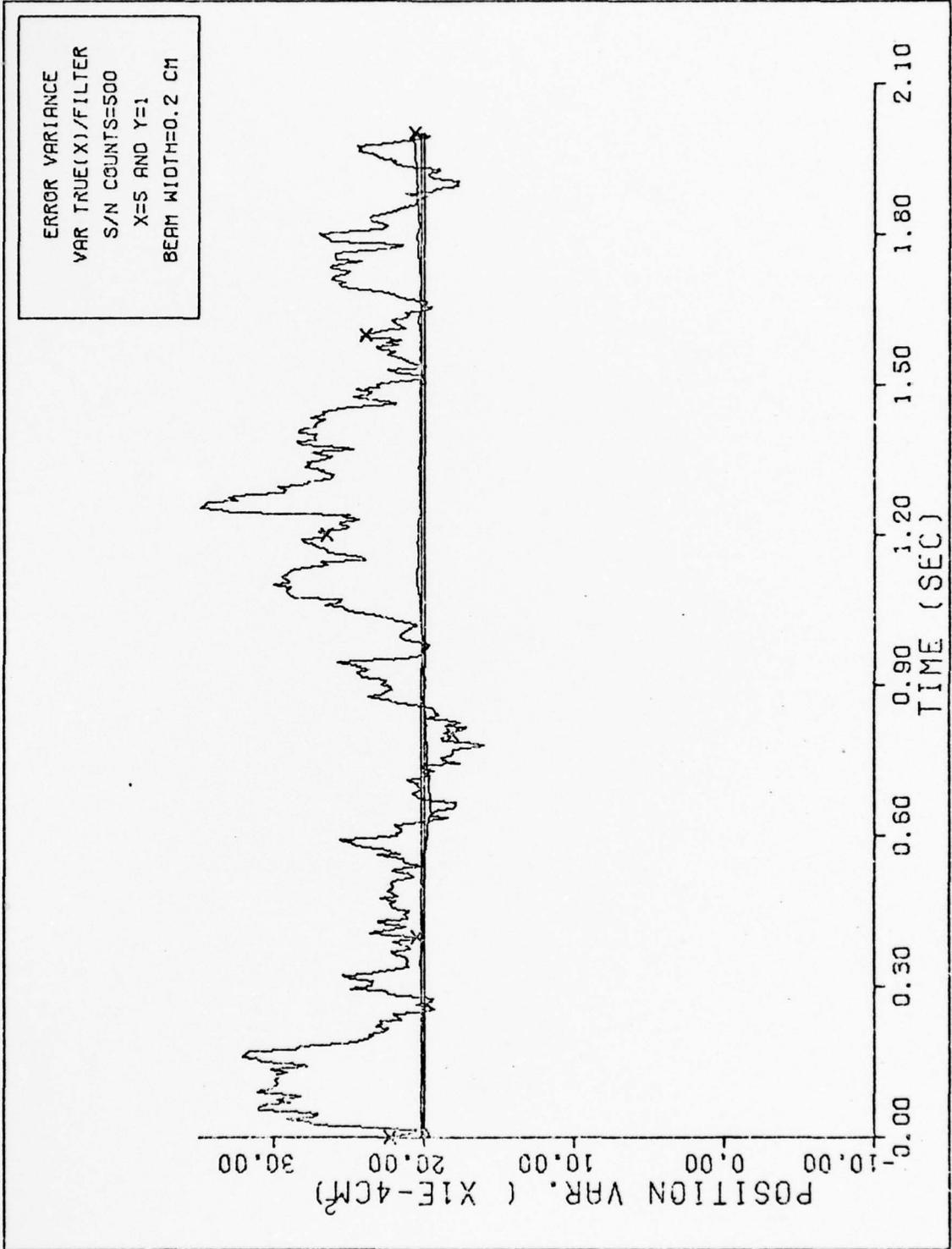
Figure 33b X-POSITION VARIANCES
(FILTER TUNING)

141
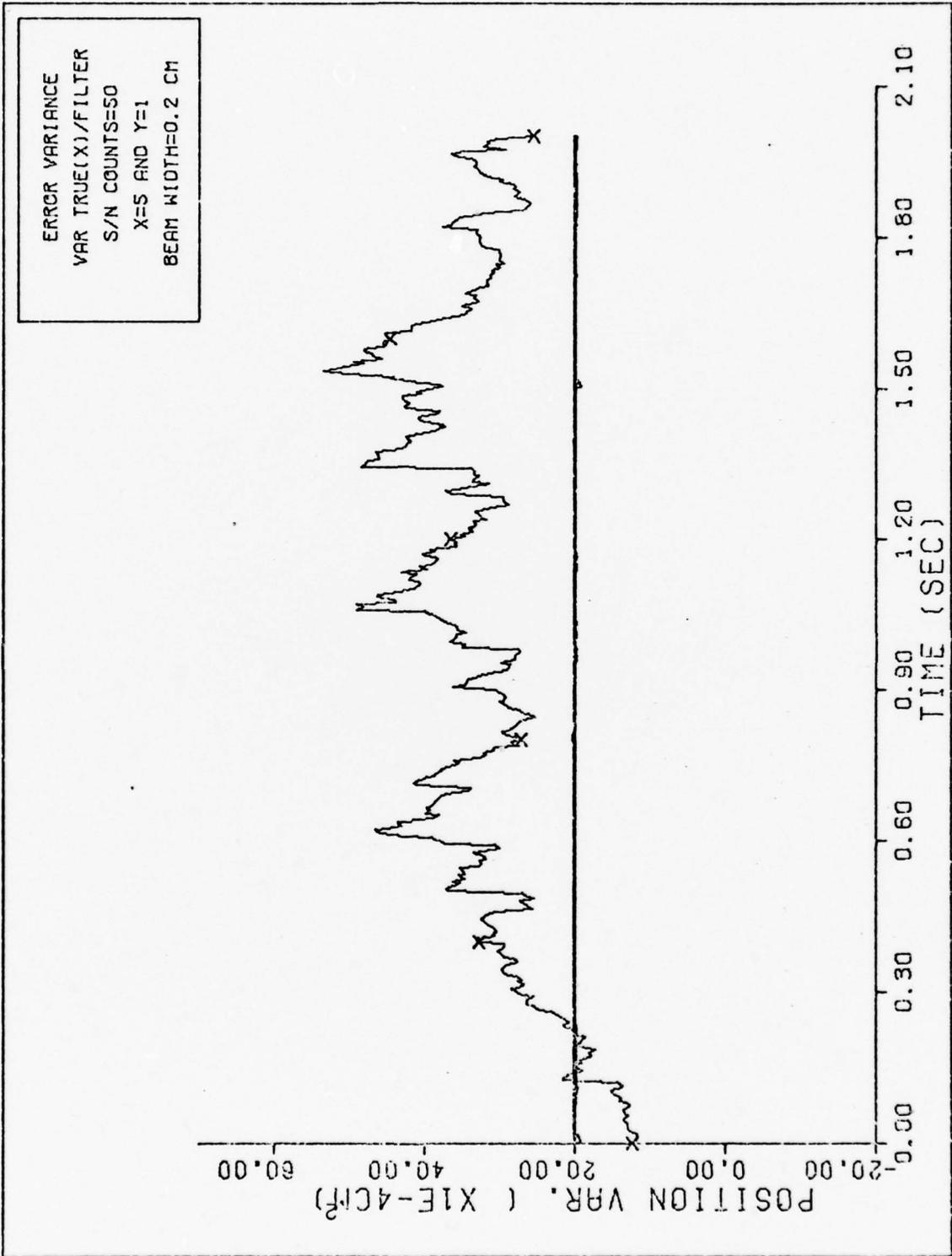
Figure 34b X-POSITION VARIANCES
(FILTER TUNING)

143
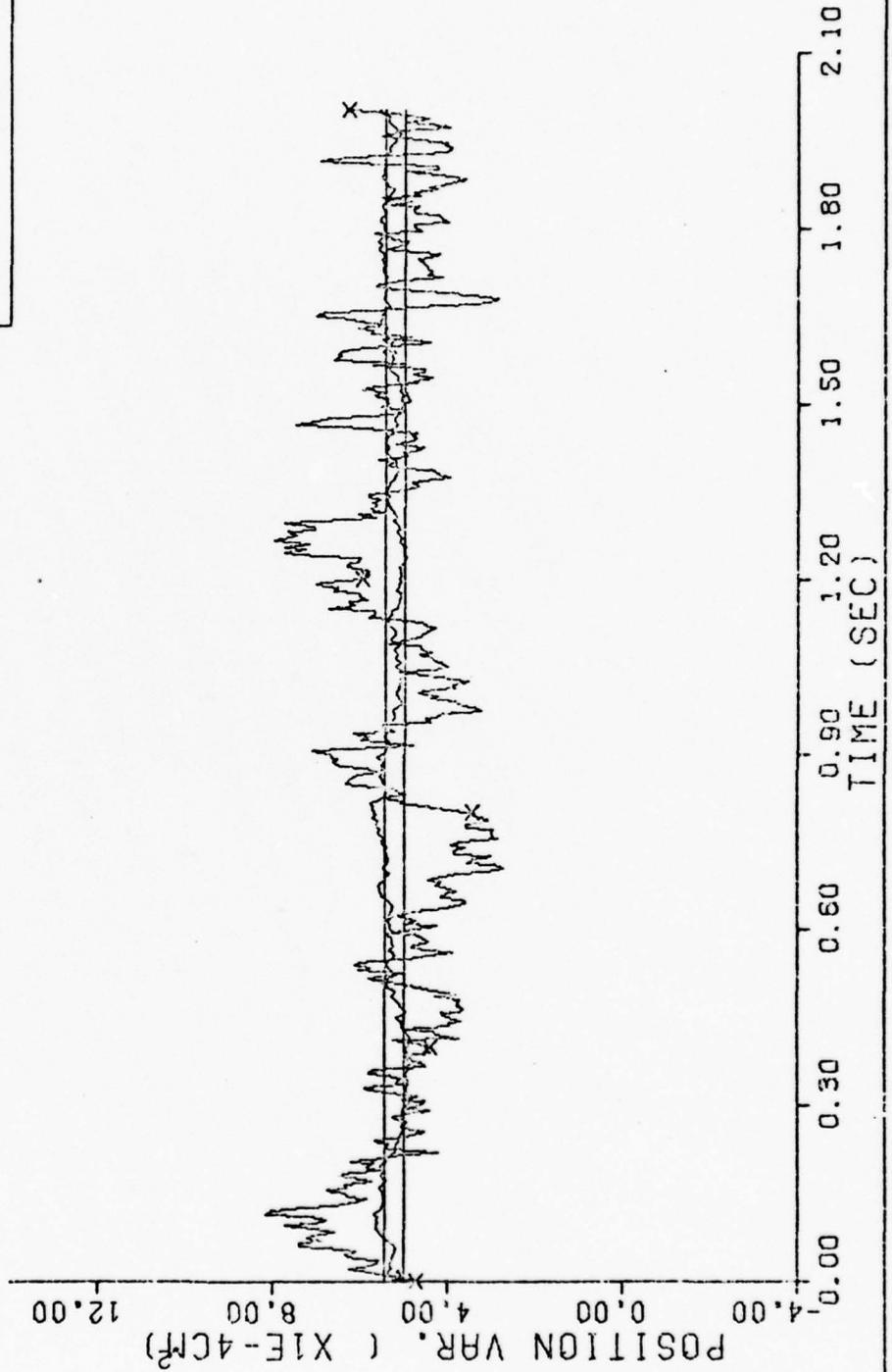
Figure 35a X-POSITION VARIANCES
(RESIDUAL MONITORING)

144
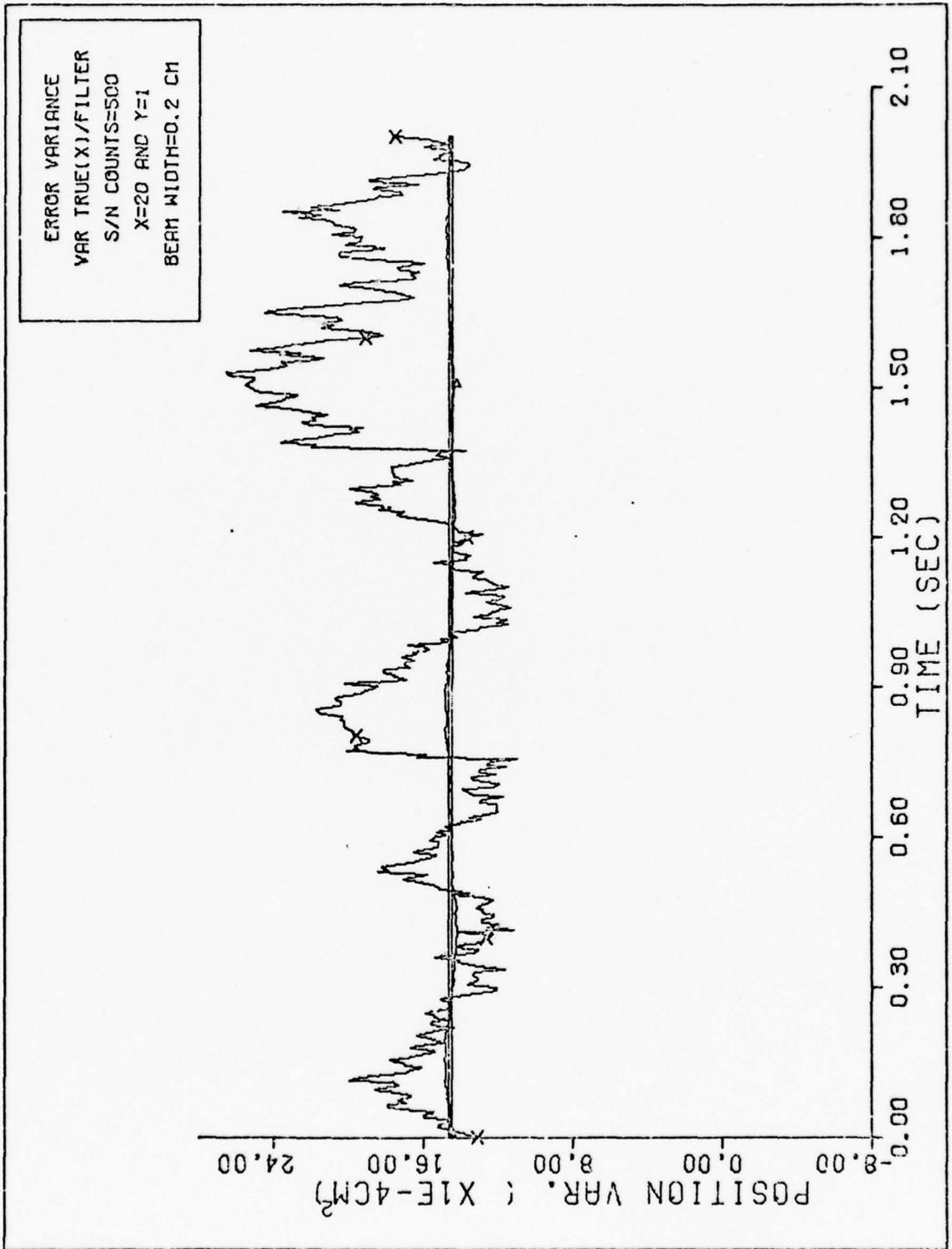
Figure 35b X-POSITION VARIANCES

(FILTER TUNING)

145
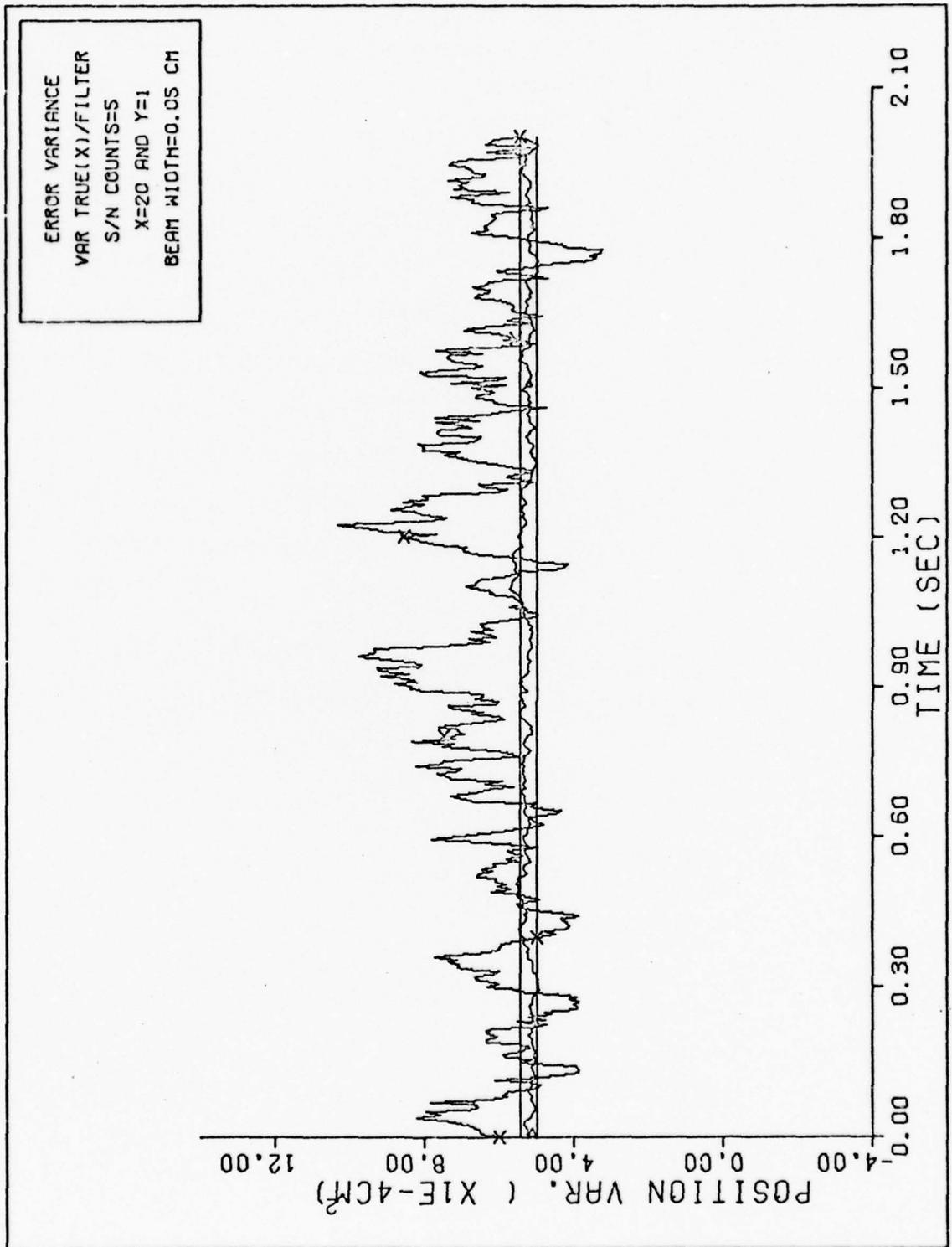
Figure 36a X-POSITION VARIANCES
(RESIDUAL MONITORING)

146
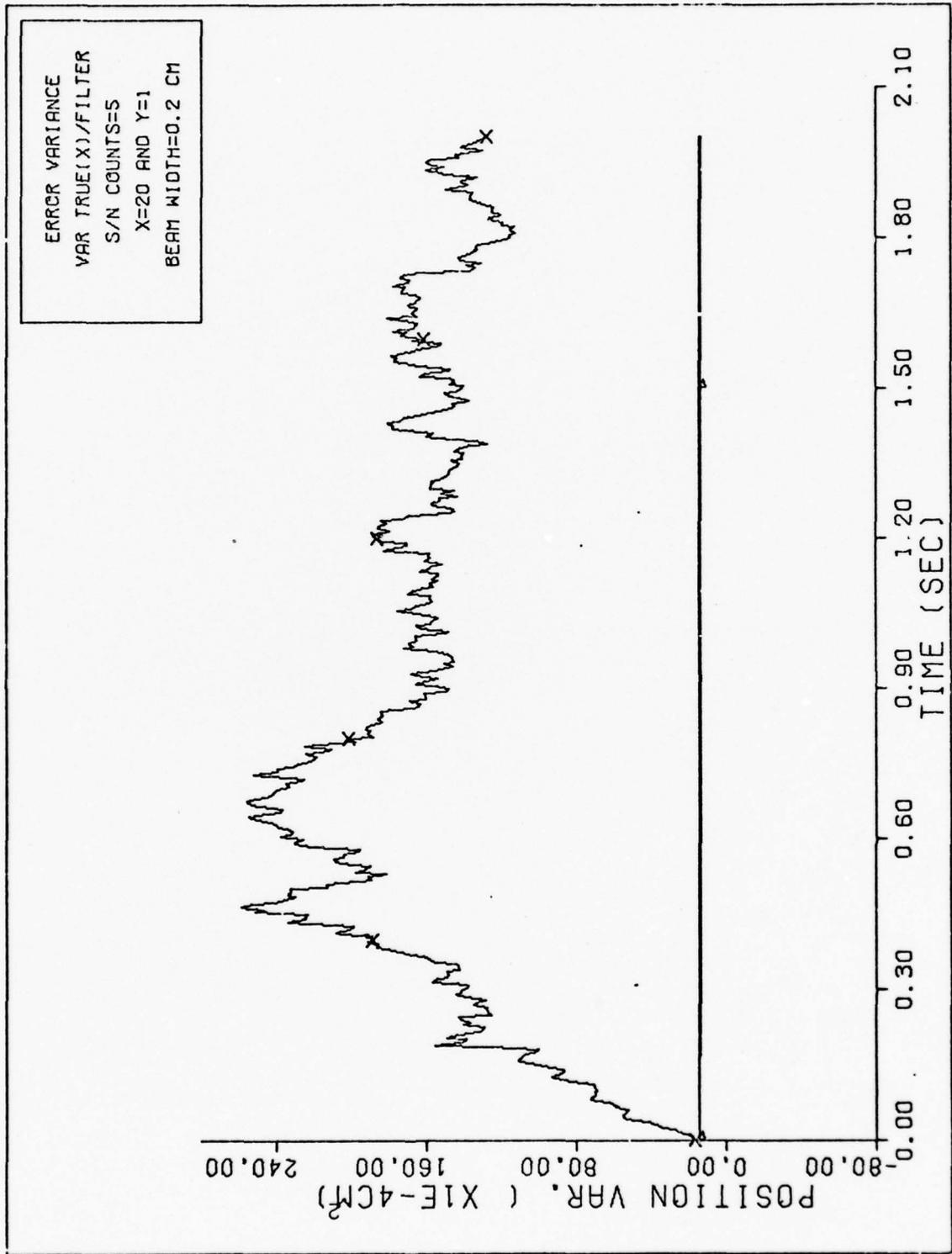
Figure 36b X-POSITION VARIANCES
(FILTER TUNING)

147

Summary

For the test cases considered, the filter's average
performance lies near the upper bound curve, which seems reasonable
since the set of parameters is selected to yield relatively
poor tracking (i.e., few signal counts per coherence time).

When dark current is included in the measurement model
of the Snyder Filter, the estimator is found to be very
sensitive to dark measurements and tracking performance is
degraded substantially.

Based on physical arguments and simulation results, residual
monitoring provides a better method to compensate for the effects
of dark current than filter tuning.

V Conclusions and Recommendations

Conclusions

Assuming no dark current, the average error performance
of the Snyder Filter can be described by steady-state upper
and lower bounds on mean square error (MSE). These bounds are
evaluated as a function of two physically motivated parameters.
These parameters are described as: (1) the average number of
signal photoelectrons observed in a coherence time of spot's
dynamics, and (2) the spot's jitter normalized by the square
of the beam width. Based on these parameters, sufficient
conditions are presented such that the bounds converge to the
actual performance. When sufficient conditions are not estab-
lished, the results from the Monte Carlo simulation indicate
that actual performance can be described by the upper bound
(in this study, the conditions correspond to low signal counts
per coherence time).

Presenting the dark measurements in the simulation gave
further insights into the tracking capabilities of the Snyder
Filter. Even for an average signal-to-noise counts ratio of
500, the results show that this estimator was highly sensitive
to dark measurements. Thus, the Snyder Filter treats all events
as signal.

Based on simulation results and physical arguments, residual
monitoring is a better way to handle the dark events rather than
increasing the beam width. This can be explained as follows:
by increasing the assumed beam width, the Snyder Filter puts less

149

weight on the dark measurements. At the same time, more events
will be assumed to be distributed from the spot's centroid.
Thus, the filter will not weigh heavily the measurement due to
the spot's Gaussian intensity profile. Also, the Snyder Filter
assumes the measurements are from a large photosensitive surface,
and therefore the beam width must be very wide to "wash" out the
dark current effects. On the other hand, residual monitoring is
a method that involves testing for dark events. In this study,
if a measurement (i.e., event location) occurred that is three
beam widths away from the position of maximum intensity, the Snyder
Filter would not process this measurement. Hence, it is unnecessary
to increase the beam width in residual monitoring. For high signal-
to-noise counts ratio, residual monitoring can be used effectively.
Results from the simulation demonstrate this physical reasoning.

Recommendations

An area of recommended research is to conduct this study in
two dimensions which would investigate different equal intensity
ellipsoidal shapes of the spot's intensity profile. Also, investi-
gation of other intensity profiles besides the assumed Gaussian-
shaped may warrant further research. The effects in filter
performance can then be investigated.

Another issue of research is how to incorporate the point-
process measurements. This study was performed in which the
rate function was chosen to have a low value, but in the real
world, rates of 10^{10} events per second can occur on the detector
surface. This imposes a very strict requirement of computation
on-line. A suggestion was made to store the photoelectrons in

charged-coupled devices. Then the Snyder Filter can process
its measurement at fixed time intervals. In this case, multiple
measurements are available to the Snyder Filter. Then, how
does one process these stored or "batch" measurements, and how
does one account for different arrival times (if at all)? This
creates an area of research that warrants further investigation.

Another illustration in how to incorporate the point-process
measurements is the issue of simultaneous events encountered in
the study. How would one process these multiple measurements
at single time instant in the structure? Obviously there is more
information concerning the spot's position of maximum intensity
(assuming no dark current). This question should be answered
with further research.

This study analyzed the Snyder Filter in open-loop configura-
tion; thus, a follow-on study with control inputs to position-
sensitive devices, such as a mirror, warrants further investiga-
tion.

A final issue and recommendation is to develop ad-hoc
methods in how to handle the dark current effects for low signal-
to-noise counts ratio. Since residual monitoring was found to
be a more effective method of compensation than filter tuning,
different methods using residual monitoring should be investigated.

Bibliography

1. Snyder, Donald L., How to Track a Swarm of Fireflies by Observing Their Flashes," IEEE Trans. on Information Theory, (Nov. 1975).

2. Gagliardi, Robert M. and Shermann Karp. Optical Communications, New York: John Wiley and Sons Inc., 1976.

3. Snyder, Donald L., Random Point Processes, New York: John Wiley and Sons Inc., 1975.

4. Rhodes, Ian B. and Donald L. Snyder, Estimation and Control Performance for Space-Time Point-Process Observations, IEEE Trans on Automatic Control (June 1977).

5. Pratt, William K., Laser Communication Systems. New York: John Wiley and Sons, Inc., 1969.

6. Yariv, Amnon., Introduction to Optical Electronics (Second Edition) Chicago: Holt; Rhinehark and Wintron, 1976.

7. Karp, Sherman, E.L. O'Neill, and R. M. Gagliardi, Communication Theory for the Free-Space Optical Channel, Proc. IEEE, 58, No. 10, 1611-1626, Oct 1976.

8. Hoverstern, E. V., R. O. Harger, and S. J. Holme, Communication Theory for the Turbulent Atmosphere, Proc IEEE, 58, No. 10 16-26, Oct 1970.

9. Maybeck, Peter S., Stochastic Estimation and Control Systems, Part 1 (Notes for EE 7.65) Wright-Patterson Air Force Base, Ohio, Air Force Institute of Technology, 1977.

10. Lewantowitz, Zdzislaw H., Laser Pointing and Tracking Using an Adaptive Extended Kalman Filter, Unpublished Thesis, Wright-Patterson Air Force Base, Ohio: Air Force Institute of Technology, July 1975.

11. Wozencraft, John M. and Irwin M. Jacobs, Principles of Communication Theory, Wiley and Sons, 1965.

12. Davenport, Wilbur B., Probability and Random Processes-An Introduction for Applied Scientists, and Engineers, McGraw-Hill Book Company, 1970.

13. Jackson, Kenneth L., Users Guide for A Generalized Monte-Carlo Analysis Program (MCAP) to Kalman Filter Design, Wright-Patterson Air Force Base, Ohio, Air Force Institute of Technology, December 1977.

Appendix A

Statistics of a Time-Space Point Process

The purpose of this appendix is to derive the statistics of a point-process in time and space based on an incremental description. In its application of this report, we define a rate function to be

$$\lambda(t,\overline{r}) = \frac{\eta}{hf_o} \int_{A_d} |U_d(t,\overline{r})|^2 d\overline{r}; \quad d\overline{r} = dxdy$$

where t denotes the time continuum, $\overline{r}$ is the dummy variable representing a location on a photodetector surface, η is the quantum efficiency, h is Planck's constant, A_d is the detector surface area, U_d is the complex field envelope, f_o is frequency of the input field, dx is differential length along x-axis, dy is differential length along y-axis and $|\cdot|$ is modulus of field envelop. This rate function describes the intensity of events appearing on the detector surface in time and space.

Shown in Fig. Al is a detector array of photon or quantum detectors which can measure the temporal and spatial location of a photon-to-electron conversion.

It is assumed that each photon-to-electron conversion or event on the detector surface is independent. Then based on its incremental properties the x-axis, y-axis, and t-axis are partitioned into intervals as shown in Fig. A-2:

Then, with $r_k \triangleq (x_k, y_k)$, the incremental description is as follows:

$$P \left[\begin{array}{l} \text{event occurring} \\ \text{in } \Delta t \Delta A \text{ at} \\ (t_k, \bar{r}_k) \end{array} \right] = \lambda(t,\bar{r}) \Delta t \Delta A + o(\Delta t \Delta A) \qquad (A-1)$$

$$P \left[\begin{array}{l} \text{no event occurring} \\ \text{in } \Delta t \Delta A \text{ at} \\ (t_k, \bar{r}_k) \end{array} \right] = 1 - \lambda(t,\bar{r}) \Delta t \Delta A + o(\Delta t \Delta A) \qquad (A-2)$$

$$P \left[\begin{array}{l} \text{more than one} \\ \text{event in } \Delta t \Delta A \\ (t_k, \bar{r}_k) \end{array} \right] = o(\Delta t \Delta A) \qquad (A-3)$$

where $P[\cdot]$ denotes probability, of $o(\Delta t \Delta A)$ are higher order terms of $o(\Delta t \Delta A)$ with $\lim\limits_{\Delta t \Delta A \to 0} \dfrac{o(\Delta t \Delta A)}{\Delta t \Delta A} = 0$. It is desirable to obtain a statistical description of event times t_k's event locations r_k's, and the number of events n, appearing on entire detector surface A_d, in a time interval $[0,t]$. Thus, we need the joint probability density function of these random variables t_k, $\bar{r}_k$, and n (i.e. $f[\{t_k\}, \{\bar{r}_k\}, n]$ where $\{\cdot\}$ denotes set of event locations or times. $f[\{t_k\}, \{\bar{r}_k\}, n]$ completely characterizes the output process from the detector surface.

To obtain $f[\{t_k\}, \{\bar{r}_k\}, n]$, consider a counting process N_t which is depicted in Fig. A-3.

The point process probability distribution for N_t (i.e., $P[N_t = n]$) will be derived using characteristic functions. Note that the process N_t is the number of events regardless of when and where the event occurred. Since N_t is a discrete random variable,

$$E[e^{jvN_t}] = e^{j(0)} P[N_t = 0] + e^{jv} P[N_t = 1]$$
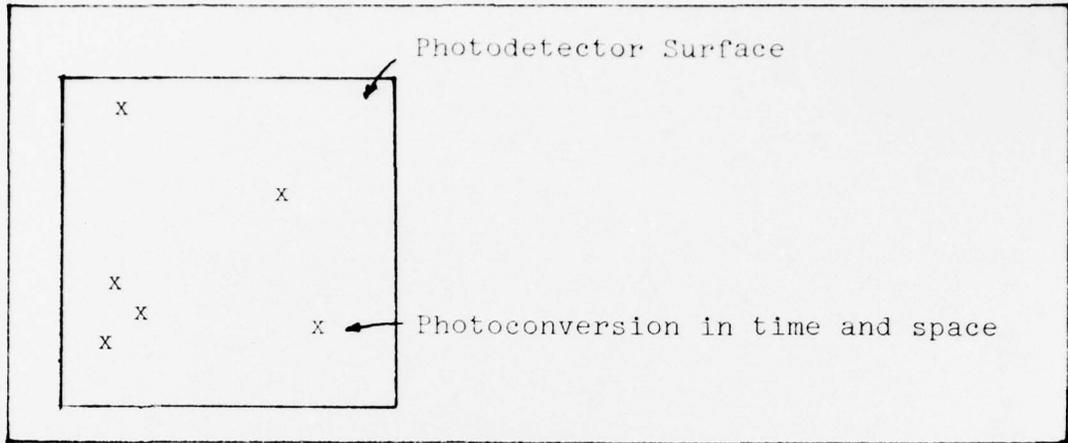$$+ e^{j2v} P[N_t = 2] + \dots + e^{jv} P[N_t = n] \qquad (A-4)$$

154

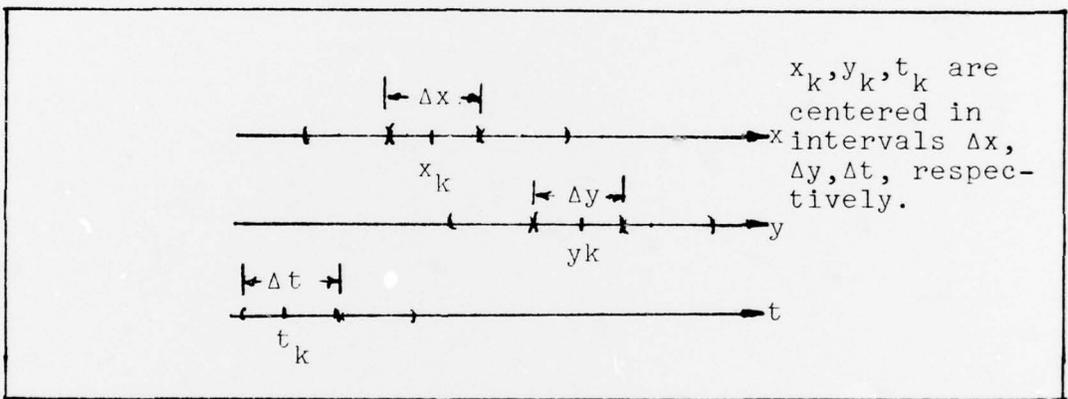Fig A.1 Infinitesimally fined detector array.

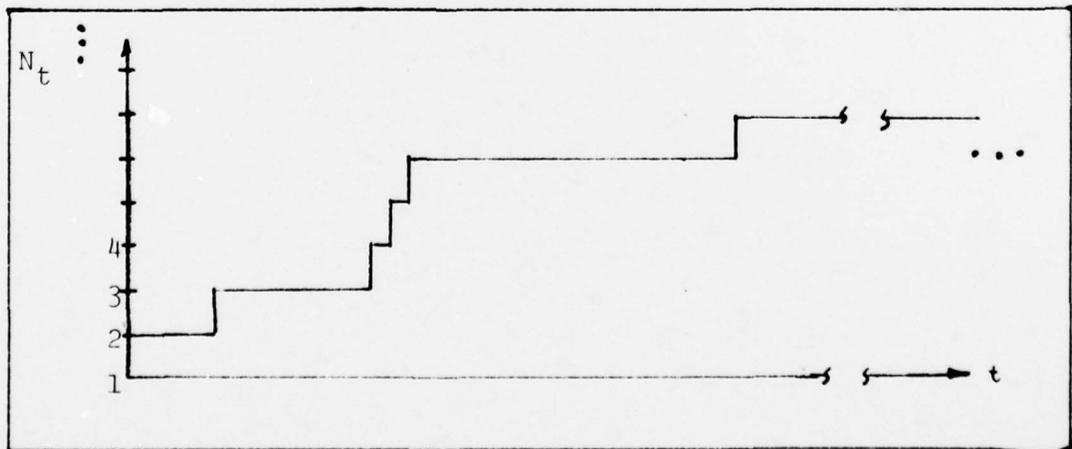Fig A.2 Partitions of x-axis, y-axis and t-axis

Fig A.3 Time-Space Counting Process

Let n_i be a random variable that is equal to a one in which
an event occurred in $\Delta t \Delta A$ or a zero in which case no event
occurred. It then follows that $N_t = \sum_{k=1}^{n} n_k$ where $k = 0,1,2,3\ldots n$.
Then Eq. (A-4) becomes

$$E[e^{jvN_t}] = E[\exp(jv \sum n_k)]$$

$$= \prod_{k=1}^{N_t} E[\exp(jvn_k)] \qquad (A-5)$$

where we have use the property of statistical independence of
events. Upon using Eqs. (A-1) through (A-3), Eq. (A-5) becomes

$$E[e^{jvN_t}] = \prod_{k=1}^{N_t} e^{jv1}\lambda(t_k,\overline{r}_k)\Delta t\Delta A + e^{jvo}[1-\lambda(t_k,\overline{r}_k)\Delta t\Delta A] + o(\Delta t\Delta A)$$

$$= \prod_{k=1}^{N_t} \lambda(t_k,\overline{r}_k)\Delta t\Delta A(e^{jv}-1) + 1 + o(\Delta t\Delta A)$$

$$= \exp \sum_{k=1}^{N_t} \ln[1 + \lambda(t_k,\overline{r}_k)\Delta t\Delta A(e^{jv}-1) + o(\Delta t\Delta t)] \quad (A-6)$$

and using $\ln(1+z) \simeq z$ for small z, Eq (A-6) becomes

$$E[e^{jvN_t}] = \exp \left[\sum_{k=1}^{N_t} \lambda(t_k,\overline{r}_k)\Delta t\Delta A(e^{jv}-1) \right]$$

Also, as $\Delta t \Delta A \to 0$, the sum in the above relation becomes a definite
integral so that

$$E[e^{jvN_t}] = \exp[(e^{jv}-1) \int_{A_d} \int_0^t \lambda(\alpha,\overline{\beta})d\alpha d\overline{\beta}] \qquad (A-7)$$

with the series expansion of $\exp(x) = 1 + x + \frac{x^2}{2!} + \cdots$

156

Eq (A-7) yields

$$E[e^{jvN_t}] = \exp(-\mu_{A_{d,t}}).$$

$$\left(1 + e^{jv}\exp(\mu_{A_{d,t}}) + \frac{e^{2jv}[\exp(\mu_{A_{d,t}})]^2}{2!} \right.$$

$$\left. + \ldots + \frac{e^{jv}[\exp(\mu_{A_{d,t}})]^n}{n!} \right) \qquad (A-8)$$

where $\quad \mu_{A_{d,t}} = \int_{A_d} \int_0^t \lambda(t,\bar{r})d\alpha, d\bar{\beta}$

Comparing Eq. (A-8) with Eq. (A-4), we obtain

$$P[N_t = n] = \frac{(\mu_{A_{d,t}})^n \exp(-\mu_{A_{d,t}})}{n!} \qquad (A-9)$$

Eq (A-9) describes the counting process N_t in interval [0,t] without regard to the event times and event locations.

To obtain $f[\{t_k\}, \{\bar{r}_k\}, n]$, and using independence of events, consider the following probabilities:

$$P[\text{one event at } (t_1, \bar{r}_1)] = \lambda(t_1, \bar{r}_1)\Delta t \Delta A \qquad (A-10)$$

$$P[\text{one event at } (t_1, r_1) \text{ and no events in } [t,T]] =$$

$$= \lambda(t_1, r_1)\Delta t \Delta A \exp\left(- \int_{A_d} \int_{t_1}^t \lambda(\alpha, \bar{\beta})d\alpha d\bar{\beta} \right. \qquad (A-11)$$

157

Finally, the probability density function of the ordered event times, event locations, and the number of events in $[0,T]$ on A_d is defined

$$f[\{t_k\},\{r_k\},n] = \lim_{\Delta t \Delta A \to 0} \frac{P[;]}{(\Delta t \Delta A)^n} \qquad (A-12)$$

where $P[;]$ denotes the probability of one event at $(t_1, \overline{r}_1)$, one at (t_2, r_2), ..., one at (t_k, r_k). Using the Eqs (A-10) and (A-11), (A-12) will yield

$$f[\{t_k\},\{r_k\},n] = \exp\left(- \int_{A_d} \int_0^t \lambda(\alpha,\overline{\beta})d\alpha d\overline{\beta}\right) \prod_{k=1}^{n} \lambda(t_k,r_k) \qquad (A-13)$$

Eq (13) completely characterizes the statistical output process of the detector surface.

Since this report simulates the detector output in one dimension, simply replace $\overline{r}_k$ with r_k, to yield

$$f\{t_k,r_k,n\} = \exp\left(- \int_\ell \int_0^t \lambda(\alpha,\beta)(\alpha,\beta)\right) \prod_{k=1}^{n} \lambda(t_k,t_k) \qquad (A-14)$$

where ℓ is detector length.

To obtain the joint probability density of event times and number of events without regard to where the events occurred (i.e. $f\{\{t_k\},n\}$), Eq (A-11) will yield the marginal density

$$f\{\{t_k\},n\} = \exp\left(- \int_\ell \int_0^t \lambda(\alpha,\beta)(\alpha,\beta)\right) \prod_{i=1}^{n} \int_\ell \lambda(\alpha,\beta)d\alpha$$

where the rate function @ $\lambda(r_k,t_k)$ is integrated over the detector length yielding no spatial dependence on $\int_\ell \lambda(\alpha_k,t_k)d\alpha$

158

This completes the statistical description of a time-space point process which characterizes the output of a photosensitive surface.

Upper and Lower Bounds on Mean Square Error

The Snyder Filter is a minimum mean-square-error estimator
that uses time-space point process observations. The purpose
of this appendix is to solve the upper and lower bounds on
mean-square-error (MSE) in order to evaluate the estimation
performance of the Snyder Filter. Also, the bounds, derived
in Ref. 4 are solved in steady-state (i.e., $\frac{dP_*}{dt} = \frac{dP^*}{dt} = 0$).
Parameters are defined from the dynamical equation of P^* and P_*
to provide meaningful physical descriptions about the actual
performance of the Snyder Filter. Based on these parameters,
sufficient conditions can be established, such that the bounds
are identical; thus, actual performance can be determined.
However, when the parameters do not establish sufficient
conditions, the bounds are no longer identical, and actual
performance is difficult to obtain. These parameters will
now be derived in one dimension which involves only estimation
performance.

The parameters are derived in one-dimension for a spot's
dynamics to be modelled as a Scalar First-Order Markov process.
In addition, this solution only involves estimation performance.

Consider the spot's centroid (or position of maximum intensity)
which is modelled by a linear stochastic differential equation

$$dx(t) = F(t)x(t)dt + G(t)d\beta(t) \qquad (B-1)$$

where $x(t) \triangleq$ position of maximum intensity, $\beta(t)$ is a Wiener
process, $F(t)$ is the state coefficient, and $G(t)$ is the noise

input coefficient. The time-space point process has a rate function modelled as

$$\lambda(r, \underset{\sim}{x}(t), t) = \mu(t)\gamma(r, \underset{\sim}{x}(t)) \qquad (B-2)$$

where

$$\mu(t) = \frac{\eta I_o(t)}{hf_o} \sqrt{2\pi}\,\sigma \qquad (B-3)$$

$$\gamma(r, x_t) = \frac{1}{\sqrt{2\pi}\sigma} \exp\{\frac{-r-x(t))^2}{2\sigma2}\} \qquad (B-4)$$

with r is a dummy variable for spatial location on detector, and σ is the beam width.

The lower bound on MSE, denoted as $P_*(t)$, is described by the following differential equation that is specialized in one-dimensional space as

$$\frac{dP_*}{dt} = F(t)P_*(t) + P_*(t)F(t) + G(t)Q(t)G(t)$$

$$- P_*^2(t)[\frac{\mu}{\sigma^2}] \qquad (B-5)$$

where $\mu = E[\mu(t)];\ E[\cdot] =$ expected value operator

 $Q(t)$ is strength of $\underset{\sim}{\beta}(t)$

The lower bound is specialized for the following assumptions:

(1) $\underset{\sim}{x}$ is modelled as an output process of a First Order Lag; then $F(t) = \frac{-1}{\tau_c}$ and its state transition function is $\phi(t,\tau) = e^{-(t-\tau)/\tau}c$. The strength Q of the white noise driving the First Order lag is given in Ref (9) as $Q = \frac{-2}{\tau_c}\sigma_x^2$

161

to obtain stationary statistics at the output. σ_x^2 is the mean square value of the spot's displacement on the detector surface (i.e. $E[\underset{\sim}{x}(t)\underset{\sim}{x}(t+\tau)] = \sigma_x^2\, e^{-(\tau)\tau_c}$) and (2) $G(t) = 1$. Eq (B-6) will yield,

$$\frac{dP_*}{dt} = \frac{-2}{\tau_c}P_* + Q - \frac{\mu}{\sigma^2} P_*^2 \qquad (B-6)$$

Solving for steady state $\dfrac{dP_*}{dt} = 0$, yields a quadratic equation which is expressed as

$$P_* = \frac{2\sigma^2}{2\tau_c \bar{\mu}} + \frac{1}{2}\left[\left(\frac{2\sigma^2}{\tau_c \mu}\right)^2 + 4\left(\frac{Q\sigma^2}{\bar{\mu}}\right)\right]^{\frac{1}{2}} \qquad (B-7)$$

$$= \left(\frac{\sigma^2}{2}\left(\frac{2}{\tau_c \bar{\mu}}\right)\right)\left(-1 + \left[1 + 4\left(\frac{Q\sigma^2}{\bar{\mu}}\right)\left(\frac{\tau_c \bar{\mu}}{2\sigma^2}\right)^2\right]^{\frac{1}{2}}\right)$$

$$= \left(\frac{\sigma^2}{2}\left(\frac{1}{\bar{\mu}\frac{\tau_c}{2}}\right)\right)\left(-1 + \left[1 + 4\left(\frac{Q\tau_c}{2\sigma^2}\right)\left(\frac{\tau_c \bar{\mu}}{2}\right)\right]^{\frac{1}{2}}\right) \qquad (B-8)$$

Dividing Eq (B-8) by σ^2 and defining $X = \dfrac{\tau_c \bar{\mu}}{2}$, $Y = \dfrac{Q\tau_c}{2\sigma^2}$ yields

$$\frac{P_*}{\sigma^2} = \frac{1}{2X}\left(-1 + [1 + 4XY]^{\frac{1}{2}}\right) \qquad (B-9)$$

But, $\bar{\mu} = \dfrac{\eta I_o}{hf_o}\sqrt{2\pi}\,\sigma$ and $Q = \dfrac{2\sigma_x^2}{\tau_c}$, and substituting into X and Y gives

$$X = \left(\frac{\eta}{hf_o} I_o \sqrt{2\pi}\sigma\right)\left(\frac{\tau_c}{2}\right) \qquad Y = \frac{\sigma_x^2}{\sigma^2}, \qquad (B-10)$$

X describes the average number of photons detected in a coherence time of the spot's dynamics and Y results in the spot's jitter normalized by the square of the beam width. These parameters are discussed in detail in Chapter II.

The relation for the upper bound on MSE is expressed as

$$\frac{dP^*}{dt} = FP^* + P^*F + GQG - \bar{\mu}P^*[P^* + \sigma^2]^{-1} P^* \qquad \text{(B-11)}$$

Applying the same assumptions as in P_*, Eq (B-11) yields

$$\frac{dP^*}{dt} = \frac{-2}{\tau_c} P^* + Q - \frac{-\bar{\mu}P^{*2}}{P^* + \sigma^2}$$

For $\frac{dP^*}{dt} = 0$ yields another quadratic equation which solves for

$$P^* = \frac{1}{2} \left\{ \frac{Q - \frac{2\sigma^2}{\tau_c}}{\frac{2}{\tau_c} + \bar{\mu}} + \left[\frac{Q - \frac{2\sigma^2}{\tau_c}}{\frac{2}{\tau_c} + \bar{\mu}} + \frac{4Q \, \sigma^2 \tau_c}{\frac{2}{\tau_c} + \bar{\mu}} - \right]^{\frac{1}{2}} \right\}$$

$$= \frac{1}{2} \left\{ \frac{Q\tau_c - 2\sigma^2}{2 + \bar{\mu}\tau_c} + \left[\frac{Q\tau_c - 2\sigma^2}{2 + \bar{\mu}\tau_c} + \frac{4Q\sigma^2\tau_c}{2 + \bar{\mu}\tau_c} \right]^{\frac{1}{2}} \right\}$$

$$= \frac{1}{2} \left\{ \frac{\sigma^2}{(1 + \frac{\bar{\mu}\tau_c}{2})} \right\} \left\{ \left(\frac{Q\tau_c}{2\sigma^2} - 1 \right) + \left[\left(\frac{Q\tau_c}{2\sigma^2} - 1 \right)^2 \right. \right.$$

$$\left. \left. + 4 \left(\frac{Q\tau_c}{2\sigma^2} \right)\left(1 + \frac{\bar{\mu}\tau_c}{2} \right) \right]^{\frac{1}{2}} \right\} \text{(B-12)}$$

Again, letting $X = \frac{\bar{\mu}\tau_c}{2}$, $Y = \frac{Q\tau_c}{2\sigma^2}$, and dividing σ^2

into Eq (B-12) yields

$$\frac{P^*}{\sigma^2} = \left(\frac{1}{2(X+1)}\right)\left((Y-1)^2 + [(Y+1)^2 + 4(X+1)Y]^{\frac{1}{2}}\right)$$

where X and Y is defined in (B-10). Note that for X >> 1 Y<<1, the bounds converge to actual mean square-error. These conditions are discussed in Chapter II.

Vita

John Mangona Santiago, Jr., was born on 15 July 1956 in Los Angeles, CA. On June 1977 he graduated from the University of California, Los Angeles with a Bachelor of Science degree, and was commissioned in the Air Force through Reserve Officer Training Corps. He entered active duty in June 1977 to be assigned to the Air Force Institute of Technology in pursuit of the Master's Degree in Electrical Engineering, Electro-optics.

Permanent Address: 2350 Elmgrove St
Los Angeles, CA 90031

165

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFIT/GEO/EE/78-4	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) FUNDAMENTAL LIMITATIONS OF OPTICAL TRACKERS		5. TYPE OF REPORT & PERIOD COVERED MS Thesis
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) John M. Santiago, Jr. 2nd Lt USAF		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Air Force Institute of Technology (ATC) AFIT/EN Wright-Patterson AFB, OH 45433		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS		12. REPORT DATE December 1978
		13. NUMBER OF PAGES 165
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) Air Force Weapons Laboratory (AFWL/ALO) Air Force Systems Command Kirtland AFB, NM		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Approved for public release; IAW AFR 190-17

JOSEPH P. HIPPS, Major, USAF
Director of Information   1-23-79

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Optical Trackers
Direct Detection
Direct Detection Optical Receiver
Time-Space Point Process
Quantum-Limit

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A problem is considered to determine the tracking capabilities of an estimator, applied in optical sensing. The estimator tracks the centroid of a one-dimensional Gaussian-shaped intensity based on time-space point process measurements. The centroid assumes to move dynamically as a First Order Gauss-Markov process. Filter performance is described by steady-state upper and lower bounds on mean-square-error (MSE) which are evaluated as a function of two physically motivated parameters: average number of photons

20. Abstract (continued)

detected in a coherence time of centroid dynamics and mean square
of the beam width. The parameters establish regions of opera-
tions where upper and lower bounds converge to the actual MSE.
Results from the Monte-Carlo simulation demonstrates the bounds
usefulness. Noise measurements from dark current or background
radiation are included in the simulation. Results show that the
filter is very sensitive to these measurements, resulting in very
poor tracking. Ad hoc methods of filter tuning and residual
monitoring are employed to improve tracking performance; results
indicate that filter performance can be improved substantially
through residual monitoring.