

State Smoothing and Filtering Using a Constant Memory for Dynamic Systems with Nonlinear Interference and Missing Observations[†]

by KERİM DEMİRBAŞ

*Department of Electrical Engineering and Computer Science (M/C 154),
University of Illinois at Chicago, P.O. Box 4348, Chicago, IL 60680, U.S.A.*

ABSTRACT: *Time-varying finite state models and interpolating functions are used to estimate the states of nonlinear discrete dynamic models with nonlinear interference and missing observations. State estimation is carried out in blocks so that its implementation requires a constant memory.*

I. Introduction

Since Kalman (1), an extensive research has been conducted on the recursive state estimation of dynamic systems under the assumption that all observations are available for all times within an observation interval. As a result, many estimation schemes have been proposed and applied for practical systems (1–9). Recently, state estimation has been considered for dynamic systems with missing observations, where the Viterbi algorithm and interpolating functions are used to estimate the states (10). But the implementation of the scheme in (10) requires an exponentially increasing memory with time. In other words, this scheme may be impractical for the state estimation for a long time interval.

In this paper, state estimation is carried out in blocks for dynamic systems with nonlinear interference and missing observations so that the proposed scheme requires a constant memory for its implementation.

II. Problem Statement

Consider the nonlinear discrete dynamic models which are defined by

$$x(k+1) = f(k, x(k), w(k)) \quad \text{the state model} \quad (1)$$

$$z(k) = g(k, x(k), I(k), v(k)) \quad \text{the observation model,} \quad (2)$$

where k indicates the discrete time; $w(k)$ and $v(k)$ are the disturbance noise and observation noise vectors at time k with zero means and known statistics; $x(0)$ and

[†] This work was carried out while the author was visiting Bilkent University, Ankara, Turkey.

$l(k)$ are the initial state vector and interference vector at time k with known statistics; $x(k)$ and $z(k)$ are the state and observation vectors at time k ; $f(\cdot)$ and $g(\cdot)$ are given functions with appropriate dimensions; and the initial state and all samples of the disturbance noise, observation noise, and interference are independent.

Let STOM be a subset of all integers in the open interval $(0, K)$, and STOA be the set of all integers in the closed interval $[0, K]$ minus STOM, where K is a big positive integer. The scheme presented below yields an estimate of the states from time 0 to time K , namely $X^K \triangleq \{x(0), x(1), x(2), \dots, x(L), x(L+1), \dots, x(2L), \dots, x(K)\}$, by assuming that only the observations at times in STOA are available (i.e. the observations at times in STOM are missing).

III. Estimation Scheme

First, the missing observations at times in STOM are estimated by using interpolating functions, and the state model is approximated by a time-varying finite state model. Then the states from time 0 to K are estimated in blocks of length L , except for the first block whose length is $L+1$. The length L is called the block length. Without loss of generality, it is assumed that $K = oL + 1$, where o is the number of blocks in $[0, K]$. The implementation of the proposed scheme requires an exponentially increasing memory with time in a block. Hence, the block length is preselected, depending upon the available memory for the state estimation.

The time-varying finite state model which approximates the state model is defined by

$$x_q(k+1) = Q(f(k, \tilde{x}(k), w_d(k))) \tag{3}$$

in which

$$\tilde{x}(k) \triangleq \begin{cases} \hat{x}(k|k) & \text{if } k = L, 2L, 3L, \dots \\ x_q(k) & \text{otherwise} \end{cases}$$

where $w_d(k)$ is a discrete disturbance noise vector which approximates the disturbance noise vector $w(k)$, and the possible values of $w_d(k)$ are denoted by $w_{d1}(k), w_{d2}(k), \dots$, and $w_{dm_q}(k)$ (2); $x_q(0)$ is a discrete initial state random vector which approximates the initial state vector $x(0)$, and the possible values of $x_q(0)$ are represented by $x_{q1}(0), x_{q2}(0), \dots$, and $x_{qm_q}(0)$, which are called the initial quantization levels; $x_q(k)$ is the quantized state at time k , whose levels are indicated by $x_{q1}(k), x_{q2}(k), \dots$, and $x_{qm_q}(k)$; $Q(\cdot)$ is the quantizer defined in (2); and $\hat{x}(k|k)$ is the estimate of the state $x(k)$ given the observation set $\{z(1), z(2), z(3), \dots, z(k)\}$, where unavailable observations are first estimated by interpolating functions.

The finite state model is represented by a trellis diagram, Fig. 1. The following metrics are assigned to all the paths of this trellis diagram. The metric of a path is the sum of the metrics of the nodes and branches along this path. The metric of a node (or quantization level) $x_{qj}(k)$ is denoted by $MN(x_{qj}(k))$ and defined as zero, except for an initial quantization level whose metric is the natural logarithm of its occurrence probability, that is

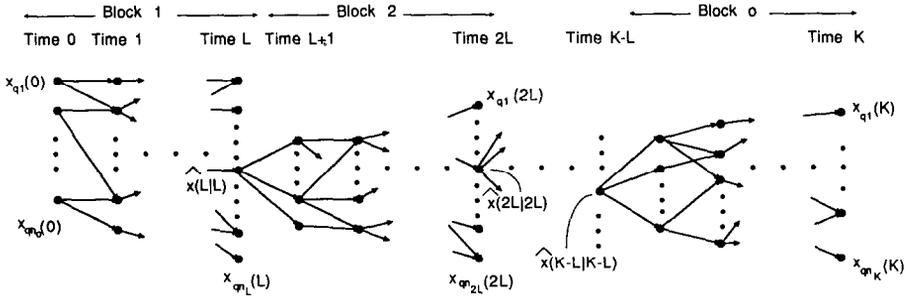


FIG. 1. Trellis diagram of state.

$$MN(x_{qi}(k)) \triangleq \begin{cases} \ln \{ \text{Prob} \{ x_q(k) = x_{qi}(k) \} \} & \text{if } k = 0 \\ 0 & \text{otherwise} \end{cases}$$

where \ln shows the natural logarithm. The metric of the branch connecting the node $x_{qi}(k - 1)$ to the node $x_{qi}(k)$ is denoted by $MB(x_{qi}(k - 1) \rightarrow x_{qi}(k))$ and defined by

$$MB(x_{qi}(k - 1) \rightarrow x_{qi}(k)) \triangleq \ln \{ T(x_{qi}(k - 1) \rightarrow x_{qi}(k)) p(z(k) | x_q(k) = x_{qi}(k)) \}$$

where $T(x_{qi}(k - 1) \rightarrow x_{qi}(k))$ is the transition probability from the node $x_{qi}(k - 1)$ to the node $x_{qi}(k)$, and $p(z(k) | x_q(k) = x_{qi}(k))$ is the conditional density function of $z(k)$ given that the quantized state $x_q(k)$ is equal to $x_{qi}(k)$. If $z(k)$ is not available, then $z(k)$ is estimated by a function which interpolates available observations in the vicinity of $z(k)$ (11). This density function is calculated by using the approximate observation model defined by

$$z(k) = g(k, x_q(k), I_d(k), v(k))$$

where $I_d(k)$ is a discrete interference vector which approximates the interference vector $I(k)$, and its possible values are denoted by $I_{d1}(k), I_{d2}(k), \dots,$ and $I_{dr_k}(k)$. Then

$$p(z(k) | x_q(k) = x_{qi}(k)) = \sum_{m=1}^{r_k} p(z(k) | x_q(k) = x_{qi}(k), I_d(k) = I_{dm}(k)) \text{Prob} \{ I_d(k) = I_{dm}(k) \},$$

where $p(z(k) | x_q(k) = x_{qi}(k), I_d(k) = I_{dm}(k))$ is the conditional density function of the observation at time k given that $x_q(k) = x_{qi}(k)$, and $I_d(k) = I_{dm}(k)$.

The state takes the quantization levels along a path through the trellis diagram from time 0 to time K . Hence, the state estimation is to find a path through the trellis diagram along which the quantization levels become the estimates of the state from time 0 to time K . It can be shown that the optimum rule which minimizes the overall error probability for choosing the correct path in a block is to choose the path with the greatest metric through the block (if there exist more than one path with the same greatest metric, to choose any one of these at random). The path with the greatest metric is chosen by the Viterbi algorithm (2, 12). This yields

the state estimates in the block. Finally, the state estimates in all blocks become the state estimates from time 0 to time K .

IV. Simulations

State estimates were obtained for many examples with white Gaussian noise and interference, and missing observations by simulations on the IBM 3081K mainframe computer. The discrete random variables given in (2) were used to approximate the initial state, disturbance noise and interference.

Simulation results of three nonlinear examples are presented in Figs 2(a)–4(c), where the first and second lines at the top left-hand corner of the figures denote the simulated state and observation models, respectively. In Figs 2(a)–4(c): $E(A)$ and $VAR(A)$ indicate the mean value and variance of the random variable A ; ACTUAL, SODS and EX. KAL. represent the actual state values, state estimates by the proposed scheme and state estimates by the extended Kalman filter which assumes zero interference; AA EOP and AA EK denote the time averaged values of the estimates by the proposed scheme and the extended Kalman filter, respectively; ER. COV. denotes the error variances of the Kalman estimates; GATE SIZE indicates the gate size used in (3); and NUM. OF DISC. FOR A indicates the number of possible values of the discrete random variable approximating the random variable A .

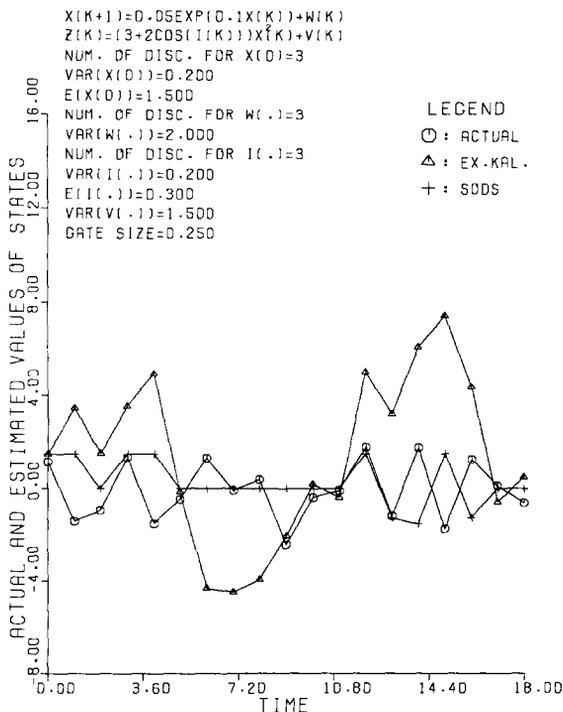


FIG. 2(a). Actual and estimated values of states.

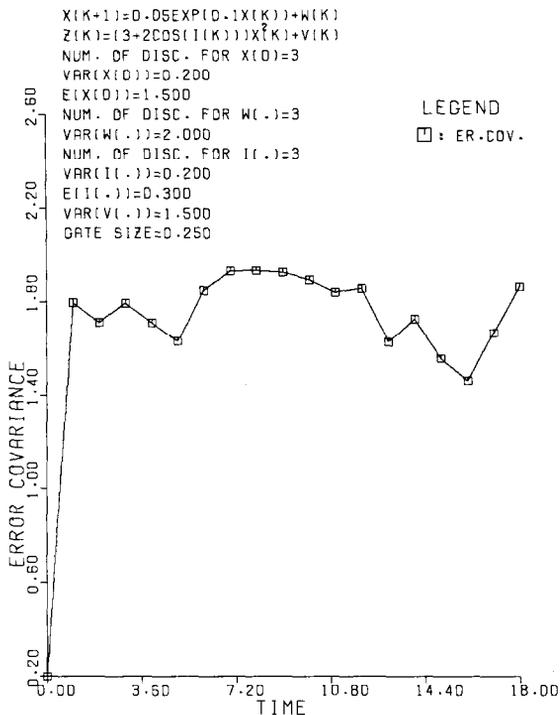


FIG. 2(b). Error variances for estimates of states.

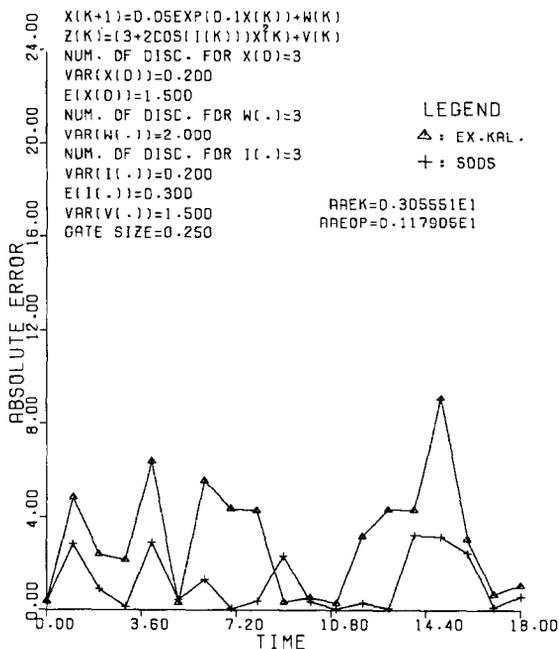


FIG. 2(c). Absolute and time-averaged absolute errors for estimates of states.

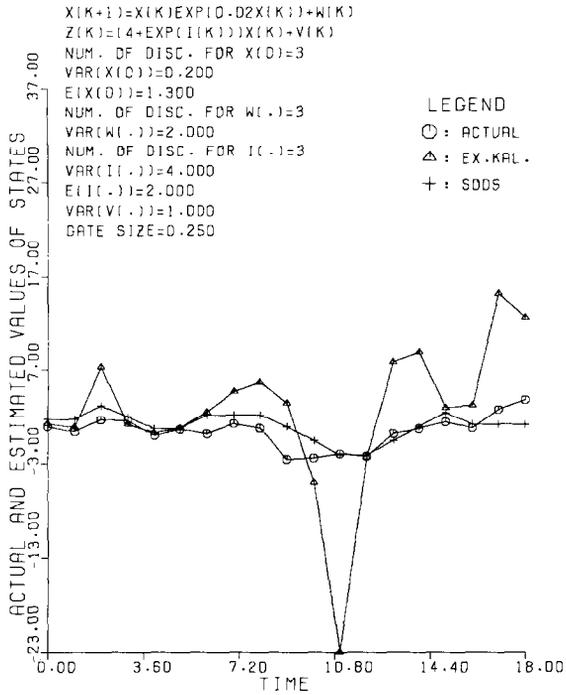


Fig. 3(a). Actual and estimated values of states.

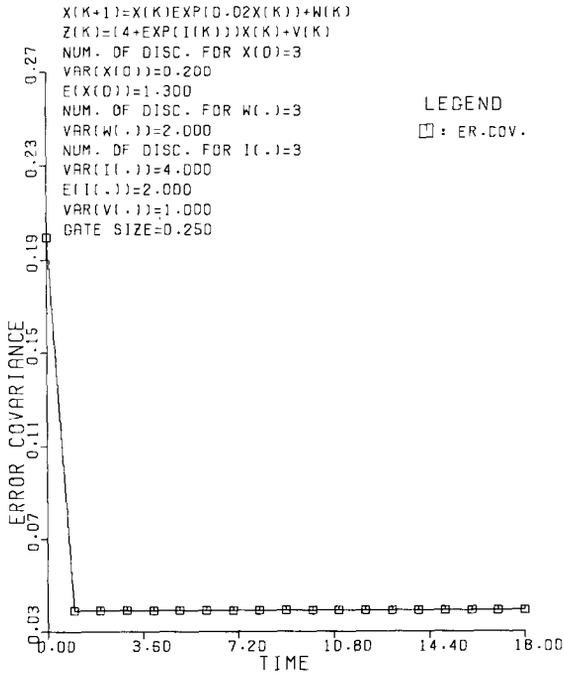


Fig. 3(b). Error variances for estimates of states.

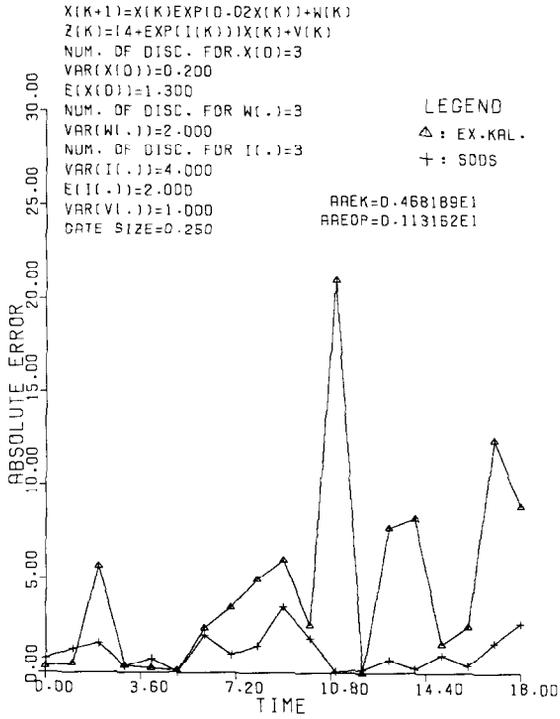


FIG. 3(c). Absolute and time-averaged absolute errors for estimates of states.

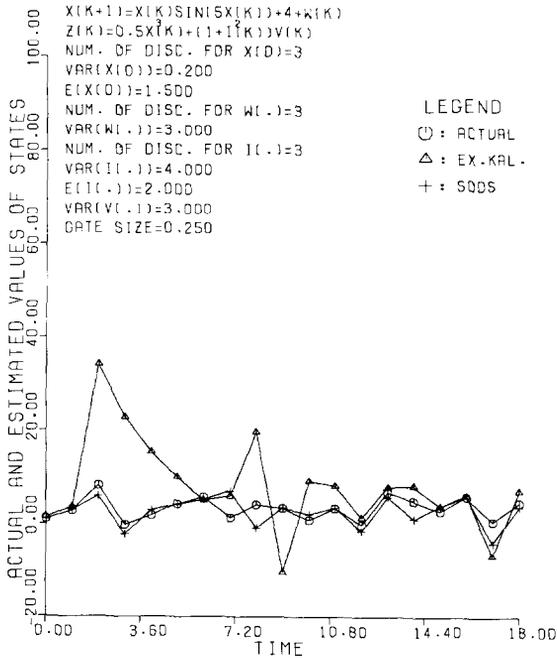


FIG. 4(a). Actual and estimated values of states.

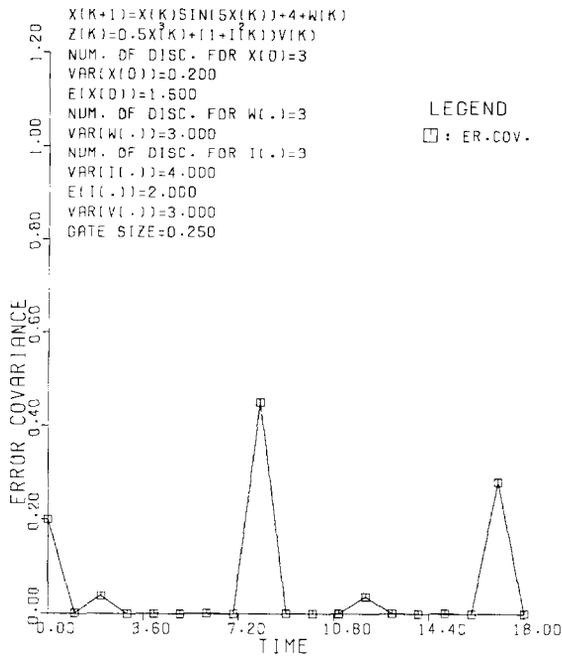


Fig. 4(b). Error variances for estimates of states.

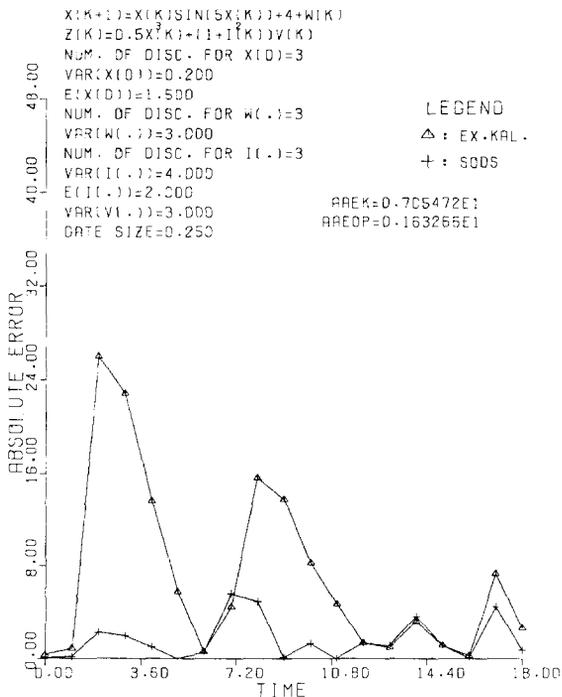


Fig. 4(c). Absolute and time-averaged absolute errors for estimates of states.

In simulations of the examples in Figs 2(a)–4(c), a block length of 6 was used and the observations at times 6, 7, 8 and 9 were assumed to be missing, namely $\text{STOM} = \{6, 7, 8, 9\}$. The observations $z(6)$, $z(7)$, $z(8)$ and $z(9)$ were estimated by using a polynomial of order 3 which interpolates the available observations $z(4)$, $z(5)$, $z(10)$ and $z(11)$. In the observation models of Figs 2(a)–3(c), a nonlinear interference is multiplied by the state, whereas a nonlinear interference is multiplied by the observation noise in the observation model of Figs 4(a)–c). The states of the models of Figs 2(a)–3(c) cannot be estimated by the extended Kalman filter because of the multiplicative nonlinear interference. On the other hand, the proposed scheme yields good estimates of the states. The divergence of state estimates by the extended Kalman filter in Figs 2(a)–4(c) is due to zero interference assumption.

V. Conclusions

An estimation scheme is proposed for nonlinear dynamic systems with nonlinear interference and missing observations. The implementation of this scheme requires a constant memory. This memory is determined by the block size selected. The accuracy of state estimates is based upon the estimation accuracy of interpolation functions used.

References

- (1) R. E. Kalman, "A new approach to linear filtering and prediction problems", *J. Basic Engng, ser. D.*, Vol. 82, pp. 35–45, 1960.
- (2) K. Demirbaş, "New smoothing algorithms for dynamic systems with or without interference", *The NATO AGARDograph Advances in the Techniques and Technology of Applications of Nonlinear Filters and Kalman Filters*, No. 256, AGARD, pp. 19–1/66, 1982.
- (3) K. Demirbaş and C. T. Leondes, "A suboptimum decoding based smoothing algorithm for dynamic systems with or without interference", *Int. J. Syst. Sci.*, Vol. 17, No. 3, pp. 499–511, 1986.
- (4) K. Demirbaş, "Maneuvering target tracking with hypothesis testing", *IEEE Trans. Aerospace Electron. Syst.*, Vol. AES-23, No. 6, pp. 757–766, 1988.
- (5) A. P. Sage and J. L. Melsa, "Estimation Theory with Applications to Communications and Control", McGraw-Hill, New York, 1971.
- (6) J. S. Medich, "A survey of data smoothing for linear and nonlinear dynamic systems", *Automatica*, Vol. 9, pp. 151–162, 1973.
- (7) T. Kailath, "A view of three decades in linear filtering theory", *IEEE Trans. Inf. Theory*, Vol. IT-20, No. 2, pp. 146–181, 1974.
- (8) R. A. Monzingo, "Discrete optimal linear smoothing for systems with uncertain observation", *IEEE Trans. Inf. Theory*, Vol. IT-21, No. 3, pp. 271–275, 1975.
- (9) C. E. Hutchinson, "The Kalman filter applied to aerospace and electronic systems", *IEEE Trans Aerospace Electron. Syst.*, Vol. AES-20, No. 4, pp. 500–504, 1984.
- (10) K. Demirbaş, "State estimation for nonlinear discrete dynamic systems with missing observations", *J. Franklin Inst.*, Vol. 327, No. 1, pp. 49–59, 1990.
- (11) J. Pachner, "Handbook of Numerical Analysis Applications", McGraw-Hill, New York, 1984.
- (12) G. D. Forney Jr, "Convolutional codes II. Maximum likelihood decoding", *Inf. Control*, Vol. 25, pp. 222–266, 1974.