Automatica, Vol. 9, pp. 163–173. Pergamon Press, 1973. Printed in Great Britain.

Discrete-Time Fixed-Lag Smoothing Algorithms*†

Algorithmes de Nivellement à Délai Fixe et à Temps Discret

Diskretzeit-Glättungsalgorithmen mit gegebener Verzögerung

Алгоритмы сглаживатия с фиксированным запаздыванием дискретного времени

JOHN B. MOORE‡

The application of Kalman filtering results yield alternative computationally stable fixed-lag smoothing algorithms including those of reduced order and minimal order.

Summary—Kalman filtering results are applied to yield alternative computationally stable fixed-lag smoothing algorithms including reduced order and minimal order fixed-lag smoothers. The reduced order smoothing algorithms are new and clearly have advantages over the more familiar algorithms. The properties of such fixed-lag smoothers are also studied.

1. INTRODUCTION

ON-LINE optimal estimation of signals in noise is of vital importance in control and communication systems. For the case of linear signal process models and additive white gaussian noise, the now widely used Kalman filter [1, 2] is the optimal on-line signal estimator. Here the term "on-line" is used in its strictest sense with the implication that the estimation of the signal is carried out simultaneously with its noisy measurement. For the case when a fixed time lag is allowed between measurement and estimation, the optimal estimator is termed a fixed-lag smoother [3–5] and consists of a Kalman filter with additional dynamics.

Of course, there are many applications, especially to control systems, for which even a small delay in computing a signal estimate is intolerable, or, at least, destroys optimality [6]. On the other hand, there are many applications, particularly to communication systems [7] and to some extent to control systems [8], where a delay sufficient to yield a useful improvement in estimation from smoothing, as opposed to filtering, is perfectly acceptable. In this connection, it is helpful to be aware that if the delay is of the order of two or three times the dominant time constant of the optimal filter, then effectively as much smoothing improvement as is possible will be achieved [9] at least, if reasonable stability properties hold for the associated filter.

It has been observed [10] that a number of fixed-lag smoothing algorithms in the literature for example, [3, 4, especially p. 234 and 5] may be normally computationally unstable and therefore impractical. To be more explicit, although the fixed-lag smoothing equations are bounded-input bounded-output stable, realisations of these in [3, 5] contain a subsystem which is unstable in the sense of Lyapunov. In [11], it is pointed out that the smoothers of [3–5] can be viewed as containing uncontrollable, unstable, blocks which can be removed without affecting the input/output properties of the smoother thus achieving a computationally stable fixed-lag smoother. In [12] a computationally stable algorithm is achieved using results developed for the more general problem of filtering in systems with multiple time delays. The results of [12] are derived in a simple manner in [13], and also in [14], using a straightforward application of Kalman filtering theory. The smoothers derived in the above references are all of dimension nN where n is the order of the signal model and N is the amount of the fixed-lag.

In this paper, we apply Kalman filtering theory, extending the ideas of [13], to derive various computationally stable fixed-lag smoothing algorithms. In particular, we consider reduced order fixed-lag smoothing structures. These structures are novel and have obvious advantages over the more familiar structures. Equations for the performance of fixed-lag smoothers are also investigated.

In order to outline the approach taken in this paper consider that noisy measurements $z(k)$ of the signal $s(k)$ are taken where $k=0, 1, 2, \ldots$. Using the notation $S_N(k)$ to denote the signal delayed by N time intervals, we have that the conditioned fixed-lag smoothed estimates $\hat{s}(k-n|k)$ are in fact

* Received 19 April 1972; revised 22 June 1972. The original version of this paper was not presented at any IFAC meeting. It was recommended for publication in revised form by Associate Editor A. Sage.

† Work supported by the Australian Research Grants Committee.

‡ Department of Electrical Engineering, University of Newcastle, New South Wales, 2308, Australia.

filtered estimates $\hat{S}_N(k|k)$ of the delayed signal $S_N(k)$. To see this we note the following relationships.

$$\begin{aligned}\hat{s}(K-N|k)&=E\{s(k-N)|z(0),\, z(1),\, \dots\, z(k)\}\\ &=E\{S_N(k)|z(0),\, z(1),\, \dots\, z(k)\}\\ &=\hat{S}_N(k|k).\end{aligned}$$

This apparently trivial observation is the basis of the approach to fixed-lag smoothing taken in this paper.

Throughout the present paper, attention is restricted to the linear–gaussian discrete-time problem; this results in the optimal fixed-lag smoother being a finite dimensional system, and straightforward conditions ensure computational stability. The fixed-lag smoothers are seen to be Kalman filters in disguise, and their derivations are simply on application of known filtering results to the usual signal process model augmented in various ways.

Extension of the results to the continuous time problem is possible, subject to one proviso. In augmenting the signal process model with delay elements, it is necessary to replace the delays by finite dimensional systems whose impulse response approximates that of an ideal delay. This results in suboptimal fixed-lag smoothers which, it should be noted, can be made arbitrarily close to optimal. This topic, together with results on nonlinear smoothing, will be studied in companion papers.

In section 2, notations are defined and Kalman filtering results are briefly reviewed. In section 3, the filtering results are applied to a signal process model augmented with delay elements to yield a fixed-lag smoother. In section 4, performance of the fixed-lag smoother is derived using the filtering equations. The results are extended in sections 5 and 6 to smoothers driven only by filtered estimates, and to reduced order smoothers, respectively. Concluding remarks are given in the final sections.

2. REVIEW OF FILTERING EQUATIONS

Consider the linear discrete signal process model

$$x(k+1)=\phi(k+1,\, k)x(k)+G(k)w(k) \tag{1}$$

$$y(k+1)=H'(k+1)x(k+1) \tag{2}$$

where $k=0,\, 1,\, \dots$ is the discrete time index. The state $x(\cdot)$ is an n-vector, the disturbance $w(\cdot)$ a p-vector, and the signal $y(\cdot)$ is an m-vector. The initial state $x(0)$ is a gaussian random vector with mean $\bar{x}_0$ and covariance P_0. The signal measurement $z(\cdot)$ is given from

$$z(k+1)=y(k+1)+v(k+1) \tag{3}$$

where the disturbances $w(\cdot)$ and $v(\cdot)$ are independent zero mean, white gaussian sequences with covariance matrices $E[w(k)w'(k)]=Q(k)$, $E[v(k)v'(k)]=R(k)$, with $R(k)$ assumed positive definite for all k, although the latter assumption can be relaxed, with subsequent increase in complexity, particularly in discussing stability.

The Kalman filter [2] with input $z(\cdot)$ and output $\hat{x}(\cdot|\cdot)$ is given from the now familiar recursive equations

$$\hat{x}(k+1|k+1)=\phi(k+1,\, k)\hat{x}(k|k)+K(k+1)\tilde{z}(k+1)$$
$$\hat{x}(0|0)=\bar{x}_0 \tag{4}$$

$$\tilde{z}(k+1)=z(k+1)-H'(k+1)\phi(k+1,k)\hat{x}(k|k). \tag{5}$$

The gain matrix $K(k+1)$ may be calculated recursively from the following equations:

$$\begin{aligned}P(k+1|k)&=\phi(k+1,\, k)P(k|k)\phi'(k+1,\, k)\\ &\quad+G(k)Q(k)G'(k)\\ P(0|0)&=P_0\end{aligned} \tag{6}$$

$$\begin{aligned}K(k+1)=P(k+1|k)H(k+1)[H'(k+1)\\ P(k+1|k)H(k+1)+R(k+1)]^{-1}\end{aligned} \tag{7}$$

$$\begin{aligned}P(k+1|k+1)&=[I-K(k+1)H'(k+1)]P(k+1|k)\\ &=P(k+1|k)[I-H(k+1)K'(k+1)].\end{aligned} \tag{8}$$

In these equations, $P(k|k)$ is the covariance of the filtering error $\tilde{x}(k|k)=x(k)-\hat{x}(k|k)$ and $P(k+1|k)$ is the covariance of the one-stage prediction error $\tilde{x}(k+1|k)=x(k+1)-\hat{x}(k+1|k)$. Recall that $\hat{x}(k+1|k)=E\{x(k+1)|Z(k)\}=\phi(k+1,k)\hat{x}(k|k)$. The signal process model and filter are indicated in Fig. 1. The symbol n-delay means the delay of an n-vector by one unit of time.

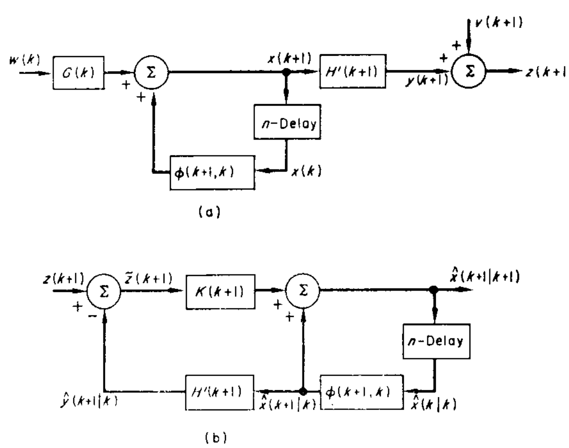

FIG. 1a. Signal process model.
1b. Kalman filter.

It proves convenient to denote the filter closed-loop system matrix by

$$F(k+1)=[I-K(k+1)H'(k+1)]\phi(k+1,\,k). \quad (9)$$

It is this matrix which determines the stability of the filter. Practical implementation of the Kalman filter, other than for brief intervals, is only possible if the filter is exponentially asymptotically stable. Sufficient, and physically reasonable, conditions guaranteeing this are known, see e.g. [15, pp. 234–243]. These are that $\phi(\cdot+1,\,\cdot)$, $\phi^{-1}(\cdot+1|\cdot)$, $G(\cdot)$, $H(\cdot)$, $Q(\cdot)$, $R(\cdot)$ and $R^{-1}(\cdot)$ are bounded, that (1) with $G(k)$ replaced by $G(k)Q^{\frac{1}{2}}(k)$ is uniformly completely controllable, and that (1) and (2) are uniformly completely observable. Henceforth in this paper, we shall assume these conditions are fulfilled.

A further consequence of these conditions, also established in [13, see p. 238], is that there exists a k_1 such that the error covariance matrices $P(k+1|k)$ and $P(k|k)$ are nonsingular for all $k \geq k_1$.

The specialization is straightforward of the Kalman filtering results to time-invariant signal process models, where ϕ, G and H are constant, and stationary signals, where Q and R are constant, and $|\lambda_i(\phi)|<1$, guaranteeing asymptotic stability of the signal process, and $k \to \infty$, or the initial time receding to $-\infty$. The notation

$$\bar{P}=\lim_{k\to\infty} P(k+1|k),$$

$$P=\lim_{k\to\infty} P(k|k)$$

and

$$K=\lim_{k\to\infty} K(k+1)$$

will be used.

3. THE FIXED-LAG SMOOTHER AS A KALMAN FILTER

The starting point for fixed-lag smoother design using the filtering results of the previous section is clearly the definition of a signal process model. For fixed-lag smoothing where the fixed-lag is N time intervals, the state to be filtered is simply the original state delayed by the amount of the fixed-lag. A reasonable signal process model is therefore simply the original signal process model (1)–(3) augmented with delay elements, see Fig. 2a. The optimal filter corresponding to this signal process model is indicated in Fig. 2b. From the diagrams above it is possible to conclude that the filter of Fig. 2b is the required fixed-lag smoother.

Before giving the equations for the smoother of Fig. 2, which are derived in [14], we observe that the (nN)-dimensional smoother of Fig. 2 has available all the fixed-lag estimates for the lag equal to $0, 1, 2, \ldots, N$ time intervals. Even though all these fixed-lag estimates are required, the smoother of Fig. 2 is frequently not of minimum dimension, as will be discussed in Section 6. For the more important case when only the smoothed estimate for the fixed-lag equal to N time-intervals is required, it becomes even easier to reduce the smoother dimension, as will be seen later.

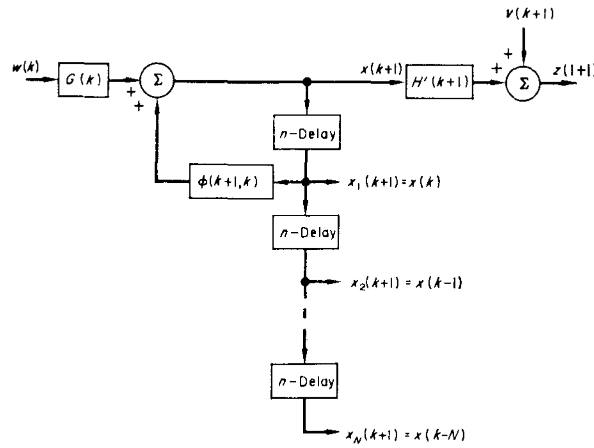

FIG. 2a. A signal process model for construction of a fixed-lag smoother.

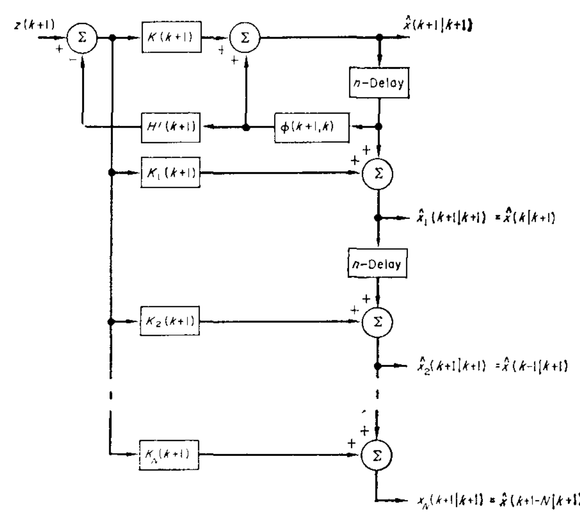

FIG. 2b. Fixed-lag smoother as a Kalman filter with dimension nN.

Returning to the details of the smoother of Fig. 2, its equations are as follows:

$$\begin{bmatrix} \hat{x}(k+1|k+1) \\ \hat{x}(k|k+1) \\ \cdot \\ \cdot \\ \cdot \\ \hat{x}(k+1-N|k+1) \end{bmatrix} = \begin{bmatrix} \phi(k+1,k) & 0 & \cdot & \cdot & \cdot & 0 \\ I_n & 0 & & & & 0 \\ 0 & I_n & \cdot & & & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & I_n & 0 \end{bmatrix} \begin{bmatrix} \hat{x}(k|k) \\ \hat{x}(k-1|k) \\ \cdot \\ \cdot \\ \cdot \\ \hat{x}(k-N|k) \end{bmatrix} + \begin{bmatrix} K(k+1) \\ K_1(k+1) \\ \cdot \\ \cdot \\ \cdot \\ K_N(k+1) \end{bmatrix} \tilde{z}(k+1) \quad (10)$$

with initial conditions $\hat{x}(0|0)=\bar{x}_0$ and $\hat{x}(-i|0)=0$ for $i\geq 0$. Figure 2b shows the structure of the fixed-lag smoother as an (nN)-dimensional system. Notice that it consists of the original filter together with an additional system, the inputs of which are the filtered estimates $\hat{x}(k|k)$ and the innovations process $\tilde{z}(k+1)$, i.e.

$$\hat{x}(k+1-i|k+1)=\hat{x}(k+1-i|k)+K_i(k+1)\tilde{z}(k+1) \quad (11)$$

for $i=1, 2, \ldots, N$.

Notice also that the only feedback in the scheme of Fig. 2b is that associated with feedback in the subsystem comprising the original filter. This means that *the smoother inherits the stability properties of the original filter*. The smoother is therefore free of the computational instability problems of the smoothers of [3–5].

We now consider the equations for the gain matrices $K_i(k+1)$ and the smoothing error covariance matrix $P(k+1-N|k+1)$, where the smoothing error is defined by

$$\tilde{x}(k+1-N|k+1)=x(k+1-N) -\hat{x}(k+1-N|k+1).$$

These are given below as

$$P_i(k+1|k)=P_{i-1}(k|k)\phi'(k+1,k)$$

$$P_i(k+1|k+1)=P_i(k+1|k)[I-H(k+1)K'(k+1)]$$

$$K_i(k+1)=P_i(k+1|k)H(k+1)[H'(k+1) P(k+1|k)H(k+1)+R(k+1)]^{-1}$$

$$P(k+1-i|k+1)=P(k+1-i|k) -P_i(k+1|k)H(k+1)K_i'(k+1)$$

for $i=1, 2, \ldots, N$. These equations are initialized by $P_i(0|0)=0$ for all $i=1, 2, \ldots, N$.

A further rewriting of these equations, making use of (7) and also (9), is instructive:

$$P_i(k+1|k)=P_{i-1}(k|k)\phi'(k+1|k) \quad (12)$$

$$P_i(k+1|k+1)=P_{i-1}(k|k)F'(k+1) \quad (13)$$

$$K_i(k+1)=P_i(k+1|k)P^{-1}(k+1|k)K(k+1) \quad (14)$$

$$P(k+1-i|k+1)=P(k+1-i|k) -P_i(k+1|k)H(k+1)K_i'(k+1) \quad (15)$$

for $i=1, 2, \ldots, N$. As written, (13) requires existence of $P^{-1}(k+1|k)$. Recall from Section 2, that the inverse certainly exists for all $k\geq$ some k_1, which may possibly be zero. If $P^{-1}(k+1|k)$ does not exist, then $P^{-1}(k+1|k)K(k+1)$ is replaced by $H(k+1)[H'(k+1)P(k+1|k)H(k+1)+R(k+1)]^{-1}$, i.e. the earlier formula for $K_i(k+1)$ is retained.

The above results may be summarized by saying that *the fixed-lag smoother as an (nN)-dimensional Kalman–Bucy filter is given by* (10) *and* (5), *see also Fig.* 2, *where the gain matrices $K_i(k+1)$ and the error covariance matrix $P(k+1-N|k+1)$ are calculated using* (12)–(15), and the original filter equations (6)–(8).

Equations (12)–(15) also make clear a further point, viz., that the computation of the gains $K_i(k+1)$ and error variances $P(k+1-i|k+1)$ of the smoother is very straightforward, having available the filter equations.

Equation (14) suggests *an alternative nN-dimensional fixed-lag smoothing structure as indicted in Fig.* 3. The equations of this smoother follow directly from (11) and (14) as

$$\hat{x}(k+1-i|k+1)=\hat{x}(k+1-i|k) +P_i(k+1|k)P^{-1}(k+1|k)K(k+1)\tilde{z}(k+1) \quad (16)$$

for $i=1, 2, \ldots, N$. Note that the $P_i(k+1|k)$ are calculated from (12) and (13) only. Again, this smoother has the same stability properties of the basic n-dimensional filter; inspection of Fig. 3 will show the absence of any feedback paths other than those associated with the production of $\hat{x}(k+1|k+1)$.

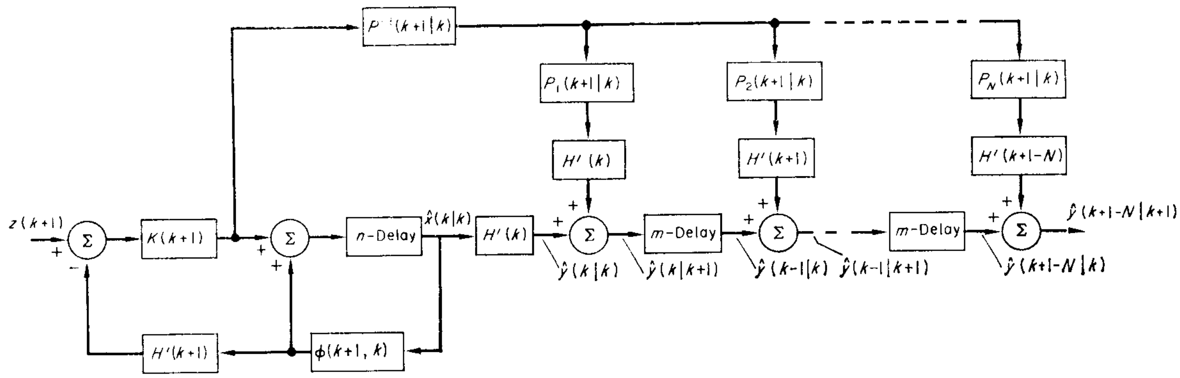

FIG. 3. Alternative fixed-lag smoothing structure with dimension nN.

Specializations of the results for the *stationary smoothing problem* are straightforward. The filter results are well known:

$$\bar{P}=\phi P\phi'+GQG' \tag{17}$$

$$K=\bar{P}H(H'\bar{P}H+R)^{-1} \tag{18}$$

$$P=\bar{P}(I-HK')=(I-KH')\bar{P}. \tag{19}$$

Equations (12)–(14) become

$$\bar{P}_i=P_{i-1}\phi'$$
$$P_i=P_{i-1}F'=\bar{P}_i(I-HK')$$
$$K_i=\bar{P}_i\bar{P}^{-1}K$$

from which one has

$$P_i=P(F')^i \tag{20}$$

$$\bar{P}_i=P(F')^i(I-HK')^{-1}=P(F')^{i-1}\phi' \tag{21}$$

and

$$K_i=P(F')^iP^{-1}K. \tag{22}$$

The form of (22) indicates that for increasing i, K_i decreases. To see this we observe that since the filter is stable, the eigenvalues of F are all less than unity and thus F^i approaches zero as i becomes large. In terms of the scheme of Fig. 2b, this means that measurements occurring at the end of a lag interval receive less weighting in computing an estimate of the state at the start of the lag interval, and a similar remark holds for other smoothing schemes discussed in subsequent sections. Of course, this phenomenon is in accord with what is intuitively reasonable. Formulae for the error covariance associated with fixed-lag estimates are discussed at some length in the next section.

4. PERFORMANCE OF THE FIXED-LAG SMOOTHER

The key property of interest in connection with smoothing is the reduction in error covariance from using smoothing rather than filtering. An explicit expression for this reduction is obtained from successive applications of (15) for $i=1, 2, \ldots, N$ as follows:

$$P(k+1-N|k+1)-P(k+1-N|k+1-N)$$
$$= -\sum_{j=k+1-N}^{k} P_{j+N-k}(j+1|j)H(j+1)K'_{j+N-k}(j+1).$$

This formula, though acceptable for computation purposes, does not explicitly make clear the non-positivity of the left hand side, i.e. the fact that smoothing yields a lower error variance than filtering. To see this reduction in error variance, we use (14) and (7) to obtain

$$P(k+1-N|k+1)-P(k+1-N|k+1-N)$$
$$= -\sum_{j=k+1-N}^{k} P_{j+N-k}(j+1|j)H(j+1)[H'(j+1)P(j+1|j)H(j+1)+R(j+1)]^{-1}\times H'(j+1)P'_{j+N-k}(j+1|j). \tag{23}$$

For the remainder of this section the stationary smoother will be considered, since the equations are simpler and the ideas involved correspond to those for the more general nonstationary case. Let us make the definition

$$D=P(F')^{-1}P^{-1}. \tag{24}$$

Then with P_s denoting the error covariance resulting with fixed-lag smoothing of N intervals, we have

$$P_s-P= -\sum_{i=1}^{N}\bar{P}_iH[H'\bar{P}H+R]^{-1}H'\bar{P}_i$$
$$= -\sum_{i=1}^{N}D^{-i}(\bar{P}-P)(D^{-i})' \tag{25}$$

after some minor manipulation. The value for $P-P_s$, the improvement obtained using smoothing as opposed to filtering, may also be obtained from the linear matrix equation

$$(P-P_s)-D^{-1}(P-P_s)(D^{-1})'=D^{-1}(\bar{P}-P)(D^{-1})' - D^{-(N+1)}(\bar{P}-P)(D^{-(N+1)})'. \quad (26)$$

Since the basic filter is asymptotically stable, the eigenvalues of F, the matrix determining its stability, must be less than 1 in magnitude. Further, since by (24), D^{-1} is similar to F' and therefore F, the eigenvalues of D^{-1} are less than 1 in magnitude. Consequently, $D^{-i}\to 0$ as $i\to\infty$, with the rate of approach to zero governed by the dominant time constant of the basic filter. After several time constants, D^{-i} will be effectively zero. We conclude that *the error covariance reduction given by* (25) *effectively attains its maximum value when the fixed lag N is of the order of two or three times the value of the dominant filter time constant.*

Of course, the error covariance reduction $P-P_s$ may not be a significant quantity relative to P, even for large N, so that before consideration is given to using a smoother rather than a filter, $P-P_s$ for large N should be compared with P. The value of (P_s-P) for large N can be obtained from (26) as the solution of the linear equation

$$D(P-P_s)D'-(P-P_s)=(\bar{P}-P) \quad (27)$$

since as noted above D^{-N} approaches zero as N becomes large.

A study of (27) for the case of a one dimensional stable signal process model $[G=h=1,\ \phi=a,\ Q=q,\ R=r]$ is of interest. For this case it can be shown that for a specified a, the maximum improvement possible from smoothing occurs when the signal to noise ratio associated with the measurements, namely $q/_r$, satisfies $q/_r=1-a^2$. The improvement for this case is $\frac{1}{2}(1-\sqrt{1-a^2})$. For the limiting case as $q/_r=1-a^2$ approaches zero, the maximum improvement possible from smoothing is 50 per cent. Another interesting result is that for the case when $N=1$, where there are no additional dynamics in the smoother, the improvement due to smoothing is maximized when $q/_r=1-a^2/_2$. The improvement for this case is $a^2/_4$. We conclude from this that the maximum possible improvement from fixed-lag smoothing where $N=1$ is 25 per cent. Note that stability implies $|a|<1$.

For signal process models of higher dimension, it is possible to prove that *for small signal to noise ratios the maximum improvement possible from smoothing is* 50 *per cent.* With large signal to noise ratios it is possible to achieve a greater improvement.

5. THE FIXED-LAG SMOOTHER AS A SYSTEM DRIVEN BY FILTERED ESTIMATES

Recall the basic filtering equation (4), which we rewrite in a slightly modified form as

$$K(k+1)\bar{z}(k+1)=\hat{x}(k+1|k+1)-\phi(k+1,k)\hat{x}(k|k).$$

This equation allows the smoothing equation (11) to be rewritten as follows, for $i=1,2,\ldots,N$:

$$\hat{x}(k+1-i|k+1)=\hat{x}(k+1-i|k) + P_i(k+1|k)P^{-1}(k+1|k)[\hat{x}(k+1|k+1) - \phi(k+1,k)\hat{x}(k|k)] \quad (28)$$

where we have also made use of (14). We infer that the smoothed estimate can be derived directly from filtered estimates. This smoothing structure is indicated in Fig. 4a as a dynamical system driven only by the filtered estimates $\hat{x}(\cdot|\cdot)$; the initial conditions are of course the same as those of the smoother of (10), as are the stability properties.

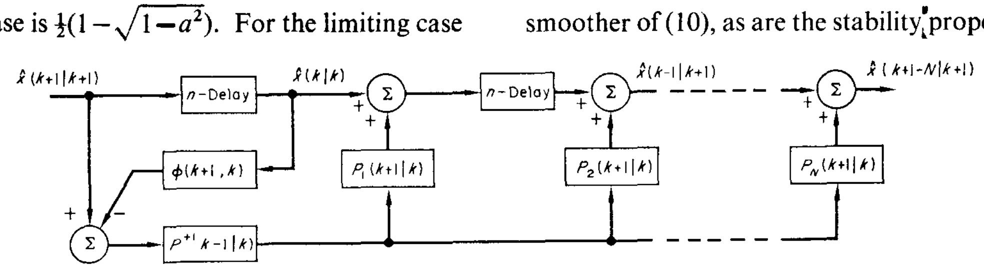

FIG. 4a. The smoother as an (nN)-dimensional system driven by filtered estimates.

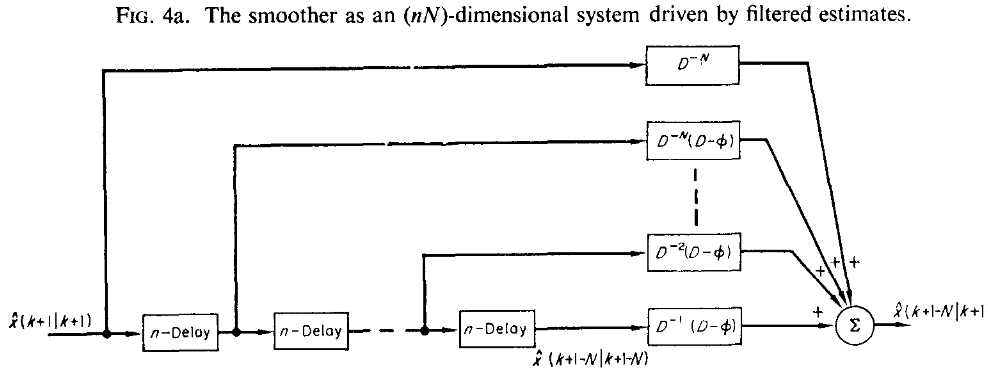

FIG. 4b. The smoother as an (nN)-dimensional system driven by filtered estimates—an alternative structure.

A variant of this arrangement can be considered. Iteration of (28) leads to

$$\hat{x}(k+1-N|k+1)=\hat{x}(k+1-N|k+1-N)$$
$$+\sum_{j=k+1-N}^{k} P_{j+N-k}(j+1|j)$$
$$\times P^{-1}(j+1|j)[\hat{x}(j+1|j+1)$$
$$-\phi(j+1|j)\hat{x}(j|j)]. \tag{29}$$

This fixed-lag smoother equation provides smoothed estimates only for a lag of N, with each smoothed estimate being a linear combination of filtered estimates, including delayed filtered estimates. The formula is a little less cluttered in the time-invariant case:

$$\hat{x}(k+1-N|k+1)=\hat{x}(k+1-N|k+1-N)$$
$$+\sum_{i=1}^{N}\bar{P}_i\bar{P}^{-1}[\hat{x}(k+1-N+i|k+1-N+i)$$
$$-\phi\hat{x}(k-N+i|k-N+i)].$$

From (19) and (21), we have $\bar{P}_i\bar{P}^{-1}=P(F')^iP^{-1}=D^{-i}$, and so the smoother equation becomes

$$\hat{x}(k+1-N|k+1)=[D^{-N}D^{-N}(D-\phi)\ \ldots\ D^{-1}(D-\phi)]\begin{bmatrix}\hat{x}(k+1|k+1)\\ \hat{x}(k|k)\\ \cdot\\ \cdot\\ \hat{x}(k+1-N|k+1-N)\end{bmatrix}. \tag{30}$$

Figure 4b illustrates the arrangement. In the time-varying case, the same structure applies, but the various gains of the gain blocks are of course varied.

6. REDUCED ORDER FIXED-LAG SMOOTHERS

The smoothers considered hitherto are not in general of minimum dimension, and we now explore techniques for deriving smoothers of lower dimension. Four preliminary comments are in order:

(1) One approach to providing smoothers of minimum dimension is to compute the impulse response of any smoother, and then use the theory of minimal realizations [16] to pass from the impulse response to a minimum dimension smoother. In the time-varying case, the computation might not be feasible, but in the time-invariant case, the Ho algorithm [16, see p. 288] could be used.

(2) For particular parameter values, it may be that the basic Kalman filter is nonminimal; this is essentially a pathological situation, and will not be considered further.

(3) When $N=1$, no additional dynamics for smoothing are required beyond those involved in the filter—see, for example, Fig. 2b. Consequently, we shall only be interested in the case $N>1$.

(4) In many applications, it may well be that *signal smoothing* rather than full state smoothing is required, i.e. smoothed estimates of the quantity $y(k)=H'(k)x(k)$ are sought.

We shall examine signal smoothing first, this being the situation where reduction in dimension of the smoother is most easily perceived. Equations (11), (16), (28) and (29) sum up the basic state smoothers hitherto developed. Since (28) and (29) are so similar, let us omit further consideration of (29). Premultiplication of the equations by $H'(k+1-i)$ yields immediately

$$\hat{y}(k+1-i|k+1)=\hat{y}(k+1-i|k)$$
$$+H'(k+1-i)K_i(k+1)\tilde{z}(k+1|k) \tag{31}$$

$$\hat{y}(k+1-i|k+1)=\hat{y}(k+1-i|k)$$
$$+H'(k+1-i)P_i(k+1|k)P^{-1}(k+1|k)$$
$$K(k+1)\tilde{z}(k+1|k) \tag{32}$$

and

$$\hat{y}(k+1-i|k+1)=\hat{y}(k+1-i|k)$$
$$+H'(k+1-i)P_i(k+1|k)P^{-1}(k+1|k)$$
$$\times[\hat{x}(k+1|k+1)-\phi(k=1,\ k)\hat{x}(k|k)]. \tag{33}$$

Figure 5 illustrates the realization of (29). With the dimension of y equal to m, the signal smoother of Fig. 5, and also the smoothers corresponding to (31) and (33), are of dimension $n+m(N-1)$. With a scalar $y(\cdot)$ for example, the increase in smoother dimension over the filter dimension is only $N-1$, while as long as $m<n$, which is usual, reduction in dimension over the earlier case occurs.

These ideas extend immediately to the problem of estimating $M(k)x(k)$, for arbitrary $M(\cdot)$; if $M(\cdot)$ has m rows, estimation of $M(k)x(k)$ with a lag of N time intervals requires a system of dimension no greater than $n+m(N-1)$.

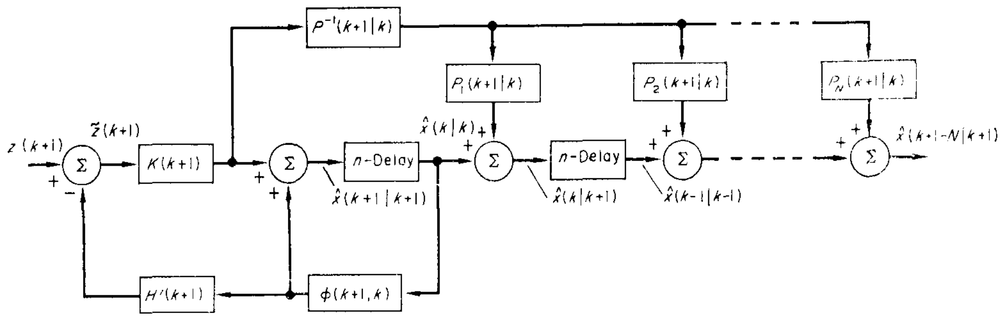

FIG. 5. Reduced order signal smoother with dimension $n+m(N-1)$.

Let us return now to the problem of state smoothing. We shall show first that *there exist state smoothers of dimension* $2n+m(N-1)$. The equations for this smoother are derived from a successive application of (11) for $i=1, 2, \ldots, N$ to yield

$$\hat{x}(k+1-N|k+1)=\hat{x}(k+1-N|k+1-N) + \sum_{j=k+1-N}^{k} K_{j+N-k}(j+1)\tilde{z}(j+1). \tag{34}$$

The filtered estimate $\hat{x}(k+1-N|k+1-N)$ is also derived in terms of the innovations $\tilde{z}(\cdot)$ as a solution of

$$\begin{aligned}\hat{x}(k+2-N|k+2-N) &= \phi(k+2-N, k+1-N)\hat{x}(k+1-N|k+1-N) \\ &\quad + K(k+2-N)\tilde{z}(k+2-N)\end{aligned} \tag{35}$$

with initial condition $\hat{x}(0|0)=0$. The implementation of this smoother is shown in Fig. 6, in which the dimension of the smoother can be observed. The dimension is greater than that of the signal smoother by n, but if $m<n$, as is usual, the smoother will have lower dimension than the state smoothers of dimension nN, at least for large enough N.

With p the number of scalar inputs to the basic signal process model, we can define smoothers of dimension $2n+p(N-1)$. These follow from the alternative signal process model of Fig. 7a, which is of dimension $2n+p(N-1)$. The point to observe is that the model is built up by delaying the inputs, rather than the states, of the basic model. The smoother is shown in Fig. 7b, and has dimension

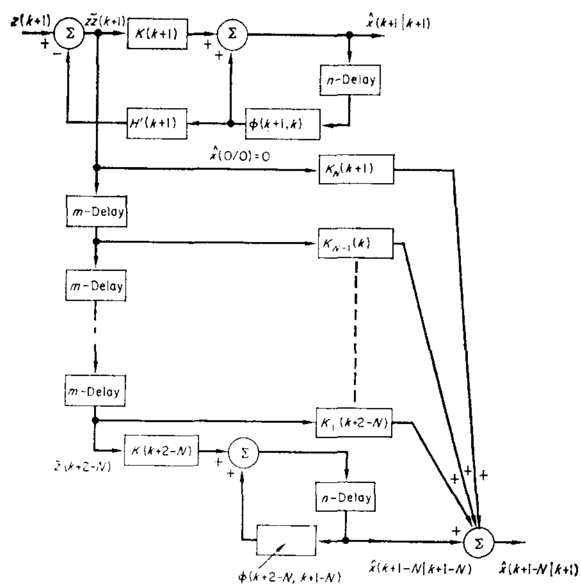

FIG. 6. Reduced order fixed-lag smoother with dimension $2n+m(N-1)$.

equal to that of the model. For the case of small p and large N, the smoother may provide simplification over those hitherto discussed. Note that calculation of the gains $M_i(\cdot)$ in Fig. 7b will be achieved by applying the usual Kalman filter equations to the system of Fig. 7a.

Finally, we shall note the sort of simplification achievable by using canonical forms for the signal process model. By way of example, we restrict attention to a single input basic signal process in the completely controllable canonical form, i.e. the state equation is

$$\begin{bmatrix} x^{(1)}(k+1) \\ x^{(2)}(k+1) \\ \cdot \\ \cdot \\ \cdot \\ x^{(n)}(k+1) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \cdot\ \cdot & 0 \\ 0 & 0 & 1 & \cdot\ \cdot & 0 \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & & 1 \\ -a_1 & -a_2 & -a_3 & \cdot\ \cdot & -a_n \end{bmatrix} \begin{bmatrix} x^{(1)}(k) \\ x^{(2)}(k) \\ \cdot \\ \cdot \\ \cdot \\ x^{(n)}(k) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{bmatrix} w(k).$$

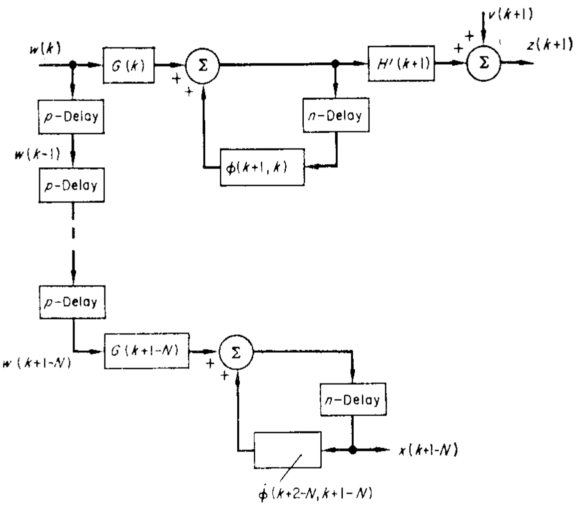

FIG. 7a. Reduced order signal process model with dimension $2n+p(N-1)$.

7. MISCELLANEOUS POINTS

The case of continuous systems with discrete measurements

It is not difficult to apply the discrete time results of this paper to this situation to obtain fixed-lag smoothers to yield smoothed state estimates at the sampling instants—the continuous system is simply replaced by a discrete system which has identical states at the sampling instants.

One point worthy of note is that as the sampling rate is increased the transition matrix ϕ approaches the unit matrix and smoothing achieves a significant error reduction as discussed in section 4.

Actually an alternative approach to this problem is used in a companion paper [20] discussing fixed-lag smoothing of signals from nonlinear systems with discrete time measurements.

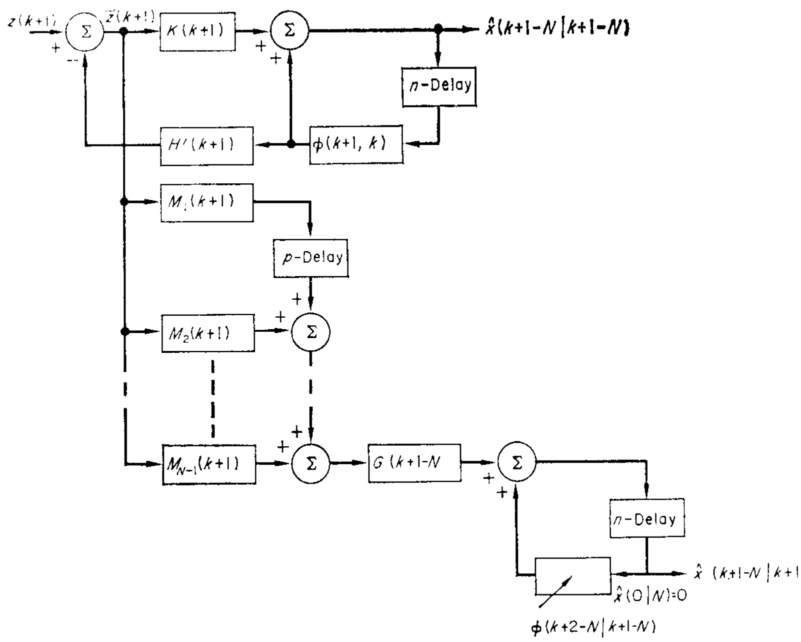

FIG. 7b. Reduced order fixed-lag smoother with dimension $2n+p(N-1)$.

It is evident that $x^{(1)}(k)=x^{(2)}(k-1)=x^{(3)}(k-2)$ $\ldots=x^{(n)}(k-n)$. Therefore, a smoothed estimate of $x^{(i)}(k-N)$ for $i>1$ is provided by a smoothed estimate of $x^{(1)}(k-N+i-1)$. It follows that it is sufficient in building a state smoother for lag N to build a smoother which computes $x^{(1)}(k-N+i-1)$ for $i=1, 2, \ldots, n$. This is a special case of signal smoothing, already discussed, with the signal dimension equal to 1. Therefore, *a state smoother of dimension* $n+(N-1)$ *can be obtained.* Simplifications along these lines can be made for multiple input systems too, but the results are complicated, depending on multivariable canonical forms.

Continuous time fixed-lag smoothing

The computational instability problem also exists for continuous-time fixed-lag smoothers [10]. In [21], a realizable, computationally stable, optimal fixed-lag smoother for continuous time signals is achieved. An appropriate duplication of the familiar, but computationally unstable smoother is required and an appropriate periodic setting to zero of the states of the Lyapunov unstable subsystems of these smoothers is also necessary.

The ideas of this paper quite naturally extend to achieve suboptimal, computationally stable, smoothers for continuous time signals. The signal model is augmented by a finite dimensional approximation to a delay and then Kalman filtering is applied to this augmented system. The resulting smoother is only optimal in the limit as the dimension of the approximation to the delay becomes infinite. The smoother in this limiting case is not realizable. Current work indicates that useful suboptimal smoothers can be built using this approach.

Fixed interval smoothing

In view of the fact that there is normally a maximum value of lag beyond which no significant improvement in smoothing occurs, there arises the possibility of using fixed lag smoothing over a fixed interval, to achieve the same estimates as fixed interval smoothing provides, without the same computational complexity that fixed interval smoothing requires because of the equivalent of a forward pass over the data, followed by a backward pass using the data or filtered estimates as discussed in [3.4]. Suppose the interval is $[0, k_1]$, and that a lag N provides all the smoothing improvement that can reasonably be achieved. Then one can obtain the equivalent of fixed interval smoothing by doing fixed-lag smoothing over an interval of length $[0, k_1+N]$ taking $H(k)=0$ or $R^{-1}(k)=0$ for $k>k^1$, with arbitrary measurements $z(k)$ for $k>k_1$.

Correlated input and output noise

It is possible to deal with the case when the input and output noise processes $w(k)$ and $v(k)$, see (1) and (3), are mutually correlated, but are still white processes. In this case, one can use an arrangement described in [17] for deriving an equivalent system with independent input and output noise. The arguments of section 3 and later sections can be applied to this new system.

Adaptive and nonlinear filtering

The fact that the smoother is nothing other than a filter means that adaptive filtering results such as those of [18] can be readily applied. Likewise, results on nonlinear filtering, see e.g. [15] and also [20], can be easily used to obtain nonlinear smoothing results.

Computational problems

For time-varying smoothing, it would be important to establish the most effective computational procedures. The equations employed may not necessarily be precisely the same as those noted in this paper. Particularly so would this be the case if the basic filter parameters were calculated by "square root filtering" methods, see e.g. [19].

Relation to existing results

The various formulas in this paper can be related to the results of [3–5] on smoothing. In view of the unsuitability of the algorithms of [3–5] on the grounds of computational instability, it seems hardly worthwhile to set out these relations here.

Further smoothing structures

It seems important to note that there are an infinity of smoothing algorithms. To grasp this point, observe that a smoother is a finite-dimensional linear system whose input is, say, the innovations process $\tilde{z}(\cdot)$ and whose output is the smoothed state estimate. It is well known that the coordinate basis of the state-space of any linear system is arbitrary. So an infinity of smoothers, not essentially different, can be obtained by varying the co-ordinate basis of the system comprising the smoother itself.

8. CONCLUSIONS

The main point to be made at this stage is the essential practicability of the smoother. It is computable via Kalman filter theory, with little increase in the complexity of the basic filtering equation. It possesses appropriate stability properties, and is realizable with any one of a number of practical structures as a finite-dimensional linear system. Fixed-lag smoothing offers virtually all the advantages of fixed interval smoothing, while it is hard to think of disadvantages.

REFERENCES

[1] R. E. KALMAN and R. S. BUCY: New results in linear filtering and prediction theory. *J. Basic Engng, Trans. Am. Soc. Mech. Engrs*, Series D, **83**, 95–108 (1961).
[2] R. E. KALMAN: A new approach to linear filtering and prediction problems. *J. Basic Engng, Trans. Am. Soc. Mech. Engrs*, Series D, **82**, 35–45 (1960).
[3] H. E. RAUCH: Solutions to the linear smoothing problem. *IEEE Trans. Aut. Control* **AC-8**, 371–372 (1963).
[4] J. S. MEDITCH: *Stochastic Optimal Linear Estimation and Control*. McGraw-Hill, New York (1969).
[5] T. KAILATH and P. FROST: An innovations approach to least-squares estimation: Part II—Linear smoothing in additive white noise. *IEEE Trans. Aut. Control* **AC-13**, 646–655 (1968).
[6] W. M. WONHAM: On the separation theorem of stochastic control. *SIAM J. Control* **6**, 312–326 (1968).
[7] D. SNYDER: *The State-Variable Approach to Continuous Estimation, with Applications to Analog Communication Theory*. M.I.T. Press, Cambridge, Mass. (1969).
[8] A. LINDQUIST: On optimal stochastic control with smoothed information. *Inform. Sciences* **1**, 55–85 (1968).
[9] B. D. O. ANDERSON: Properties of optimal linear smoothing. *IEEE Trans. Aut. Control* **AC-14**, 114–115 (1969).
[10] C. N. KELLY and B. D. O. ANDERSON: On the stability of fixed-lag smoothing algorithms. *J. Franklin Inst.* (to appear).
[11] S. CHIRARATTANANON and B. D. O. ANDERSON: The fixed-lag smoother as a stable finite-dimensional linear system. *Automatica* **7**, 657–669 (1971).
[12] R. PREMIER and A. G. VACROUX: On smoothing in linear discrete systems with time delays. *Int. J. Control* **13**, 299–303 (1971).

[13] M. FAROOQ and A. K. MAHALANABIS: A note on the maximum likelihood state estimation of linear discrete systems with multiple time delays. *IEEE Trans. Aut. Control* **AC-16,** 105–106 (1971).

[14] K. K. BISWAS and A. K. MAHALANABIS: A Recursive Algorithm for Fixed Lag Smoothing. *Symposium on Networks, Systems and Controls*, Indian Institute of Technology (1970).

[15] A. H. JAZWINSKI: *Stochastic Processes and Filtering Theory*. Academic Press, New York (1970).

[16] R. E. KALMAN, P. FALB and M. A. ARBIB: *Topics in Mathematical System Theory*. McGraw-Hill, New York (1968).

[17] R. E. KALMAN: *New Methods in Wiener Filtering Theory. Proceedings Symposium Eng. Appl. Random Function Theory and Probability* (Edited by J. L. BOGDANOFF and F. KOZIN). Wiley, New York (1963).

[18] R. K. MEHRA: On the identification of variances and adaptive Kalman filtering. *IEEE Trans. Aut. Control* **AC-15,** 175–184 (1970).

[19] A. ANDREWS: A square root formulation of the Kalman covariance equation. *AIAA J.* **6,** 1163–1166 (1968).

[20] J. B. MOORE and P. TAM: Fixed-lag smoothing of nonlinear systems with discrete measurements. *Inform. Sciences* (Also *Proceedings Fifth Hawaii International Conference on System Sciences*) (1972).

[21] P. K. TAM, J. B. MOORE and B. D. O. ANDERSON: Stable realizations of the optimal continuous-time fixed by smoothingequations. Proceedings of third symposium on nonlinear estimation theory and its applications (1972).

Résumé—Les résultats de filtrage de Kalman sont appliqués pour donner des algorithemes alternatifs de nivellement à délai fixe et de calcul stable, y compris des niveleurs à délai fixe d'ordre réduit et d'ordre minimal. Les algorithmes de nivellement d'ordre réduit sont nouveaux et ont des avantages distincts sur les algorithmes plus familiers. Les propriétés de tels niveleurs à délai fixe sont également étudiées.

Zusammenfassung—Resultate der Kalman-Filterung wurden angewandt, um alternativ berechenbare stabile Glättungsalgorithmen mit gegebener Verzögerung einschließlich Glättungseinrichtungen reduzierter und minimaler Ordnung mit gegebener Verzögerung zu liefern.

Die Glättungsalgorithmen reduzierter Ordung sind neu und haben gegenüber den bekannteren klare Vorzüge. Die Eigenschaften solcher Glättungseinrichtungen mit gegebener Verzögerung werden ebenfalls studiert.

Резюме—Результаты фильтрования Кальмана применяются для получения альтернативных устойчивых в вычислениях алгоритмов сглаживания с фиксированным запаздыванием, включая сглаживатели с фиксированным запаздыванием сокращенного и минимального порядка. Алгоритмы сглаживания сокращенного порядка являются новыми и очевидно имеют преимущества перед более знакомыми алгоритмами. Рассматриваются также свойства сглаживателей с фиксированным запаздыванием.