Now define

$$Y_i=\begin{bmatrix} y(-i) & y(-1-i) & \cdots & y(1-p-i) \\ y(1-i) & y(-i) & & y(2-p-i) \\ \cdot & & & \\ \cdot & & & \\ y(N-1-i) & y(N-2-i) & \cdots & y(N-p-i) \end{bmatrix}$$

$$U_i=\begin{bmatrix} u(-i) & u(-1-i) & \cdots & u(1-p-i) \\ u(1-i) & u(-i) & & u(2-p-i) \\ \cdot & & & \\ \cdot & & & \\ u(N-1-i) & u(N-2-i) & \cdots & u(N-p-i) \end{bmatrix}.$$

Substituting $\epsilon(k)$ into Ω, we can easily verify that

$$\Omega=-Y_0-\sum_{i=1}^{p} a_iY_i+\sum_{i=0}^{p} b_iU_i.$$

REFERENCES

[1] K. I. Åström and P. Eykhoff, "System identification—A survey," *Automatica*, vol. 7, pp. 123–162, 1971.
[2] D. W. Clarke, "Generalized-least-squares estimation of the parameters of a dynamic model," in *Proc. of the IFAC Symp. Identification in Automatic Control Systems*, 1967, paper 3.17.
[3] T. C. Hsia, "New generalized least squares estimation algorithm for system identification," in *Proc. IEEE Conf. Decision and Control*, 1973, pp. 755–759.
[4] J. L. Talmon and A. J. W. Van der Boom, "On the estimation of the transfer function parameter of process and noise dynamics using a single stage estimator," in *Proc. 3rd IFAC Symp. Identification and System Parameter Estimation*, 1973, pp. 711–720.
[5] A. S. Goldberger, *Econometric Theory*. New York: Wiley, 1964.
[6] T. C. Hsia, "A new iterative least squares algorithm for system parameter identification," Dep. Elec. Eng., Univ. California, Davis, Res. Rep. 132, 1974.

A Generalized Likelihood Ratio Approach to the Detection and Estimation of Jumps in Linear Systems

ALAN S. WILLSKY, MEMBER, IEEE, AND HAROLD L. JONES, MEMBER, IEEE

***Abstract*—We consider a class of stochastic linear systems that are subject to jumps of unknown magnitudes in the state variables occurring at unknown times. This model can be used when considering such problems as estimation for systems subject to possible component failures and the tracking of vehicles capable of abrupt maneuvers. Using Kalman–Bucy filtering and generalized likelihood ratio techniques, we devise an adaptive filtering system for the detection and estimation of the jumps. An example that illustrates the dynamical properties of our filtering scheme is discusssed in detail.**

I. INTRODUCTION

In recent years the Kalman–Bucy filter has been applied to a wide variety of practical problems. In some cases the dynamical system being studied is linear and can be modeled quite accurately. For such problems the Kalman–Bucy filter performs extremely well. However, there are many applications for which standard Kalman–Bucy filtering techniques are inadequate and "adaptive filtering" techniques are required [1]–[18].

Manuscript received April 1, 1974; revised September 8, 1975. Paper recommended by H. W. Sorenson, Chairman of the IEEE S-CS Estimation and Identification Committee. This work was developed at The Analytic Sciences Corporation, Reading, MA, under USAF Contract FO4701-74-C-0095.

A. S. Willsky is with the Decision and Control Sciences Group, Electronic Systems Laboratory, Department of Electrical Engineering, Massachusetts Institute of Technology, Cambridge, MA.

H. L. Jones is with The Analytic Sciences Corporation, Reading, MA.

In this short paper we consider an adaptive filtering problem for linear systems subject to abrupt changes. As discussed in [1] and [22], such models can be used to study systems subject to system component failures, systems involving small nonlinearities, and systems for which we wish to use a filter of lower dimension than the actual system state (e.g., the design of a tracking filter based on a constant velocity model, ignoring possible vehicle accelerations [4]).

Our approach, which is a modification and generalization of the techniques in [4] and [16], is based on the assumption that abrupt system changes may occur but that they occur infrequently, i.e., that our basic model is correct except for sporadic system anomalies, such as failures. Given this assumption, the philosophy of our approach is as follows: we implement a Kalman–Bucy filter based on the assumption of no abrupt system changes, and we design a secondary system that monitors the measurement residuals of the filter to determine if a change has occurred and adjusts the filter accordingly. The reasoning behind this structure is that, since changes occur relatively infrequently, we do not wish to degrade the performance of our filter under normal conditions by requiring our state estimator to be directly sensitive to system changes.

As this work was primarily motivated by the problem of failure detection, we have concentrated a great deal of our attention on the problem of detection of jumps; however, our detection method, which is based on the generalized likelihood ratio [20], leads directly to two filter compensation methods described in Section IV. We refer the reader to [3], [4], [9], [10], [16], and [18] for other adaptive filtering techniques involving the use of detection-theoretic notions.

II. LINEAR STOCHASTIC SYSTEMS WITH UNKNOWN JUMPS

Consider the following discrete-time dynamical system (considered also in [16]):

$$x(k+1)=\Phi(k+1,k)x(k)+\Gamma(k)w(k)+\delta_{\theta,k+1}\nu \tag{1}$$

$$z(k+1)=H(k+1)x(k+1)+v(k+1) \tag{2}$$

where $x(k)\in R^n$ is the state, with Gaussian initial condition $x(0)$ with mean $\hat{x}_0$ and covariance P_0. In addition, $z\in R^p$ is the observation, and $\{w(k)\}$ and $\{v(k)\}$ are independent, zero mean, white Gaussian sequences with $E[w(k)w(k)']=Q(k)$ and $E[v(k)v(k)']=R(k)>0$. The term $\delta_{\theta,k+1}\nu$ represents a possible jump in one or more of the state variables. Here θ is an unknown positive integer, which assumes a finite value if a jump occurs and takes the value $+\infty$ if there is no jump. Also δ_{ij} is the Kronecker delta and ν is the unknown size of the random jump. We either assume that ν is completely free or that there are a finite number of possible "jump directions" $f_1,\cdots,f_N$ with $\nu=\alpha f_i$ for some unknown i and unknown scalar α.

In the next section we develop a technique for the detection and estimation of such jumps. Also, our method can be used to detect multiple jumps by separate detection of and compensation for individual jumps. As discussed in [22], this model can be used to consider such problems as actuator, sensor, and plant changes and failures and the detection of higher order, unmodeled effects (such as acceleration in the constant velocity model mentioned in Section I) by state augmentation. In this manner we can consider step and ramp-type phenomena.

III. THE GENERAL LIKELIHOOD RATIO TECHNIQUE

Consider the system (1), (2). We wish to design an adaptive filtering system for the estimation of the state $x(k)$. Based on the comments in Section I, we assume that we have implemented a Kalman–Bucy filter based on the "no-jump" ($\theta=\infty$) hypothesis H_0.

$$\hat{x}(k+1|k)=\Phi(k+1,k)\hat{x}(k|k) \tag{3}$$

$$\hat{x}(k|k)=\hat{x}(k|k-1)+K(k)\gamma(k) \tag{4}$$

where $\gamma(k)$ is the measurement residual

$$\gamma(k)=z(k)-H(k)\hat{x}(k|k-1) \tag{5}$$

and the gain, error covariance, and residual covariance satisfy

$$K(k)=P(k|k-1)H'(k)V^{-1}(k) \tag{6}$$

$$P(k+1|k)=\Phi(k+1,k)P(k|k)\Phi'(k+1,k)+\Gamma(k)Q(k)\Gamma'(k) \quad (7)$$

$$P(k|k)=P(k|k-1)-K(k)H(k)P(k|k-1) \quad (8)$$

$$V(k)=H(k)P(k|k-1)H'(k)+R(k). \quad (9)$$

Straightforward calculations [22] allow us to express the residual as a sum of two terms

$$\gamma(k)=G(k;\theta)\nu+\gamma_1(k) \quad (10)$$

where γ_1 is a zero mean white noise sequence with covariance $V(k)$ (it represents the actual measurement residual if a jump does not occur) and G can be precomputed (see the Appendix for the equations defining G). We can then perform a generalized likelihood ratio (GLR) [20] test to decide if a jump has occurred (the H_1 hypothesis). The details of this procedure are described in [22]. Essentially, we compute the maximum likelihood estimates (MLE's) $\hat{\theta}(k)$ and $\hat{\nu}(k)$ based on $\gamma(1),\cdots,\gamma(k)$ and the hypothesis H_1. These values are then used in computing the log-likelihood ratio $l(k)$ for H_1 versus H_0, given the observed residuals $\gamma(1),\cdots,\gamma(k)$. Using the fact that all the relevant densities are Gaussian, we have

$$\hat{\nu}(k)=C^{-1}[k;\hat{\theta}(k)]d[k;\hat{\theta}(k)] \quad (11)$$

where C is deterministic

$$C(k;\theta)=\sum_{j=\theta}^{k} G'(j;\theta)V^{-1}(j)G(j;\theta) \quad (12)$$

and d is a linear combination of the residuals

$$d(k;\theta)=\sum_{j=\theta}^{k} G'(j;\theta)V^{-1}(j)\gamma(j). \quad (13)$$

The operation on the residuals in (13) can be interpreted as a matched filter (MF) or a least squares estimate of the jump ν assuming that θ is known and that we have no *a priori* information about the value of ν. In this case $C^{-1}(k;\theta)$ is the error covariance of our estimate of ν.

The MLE $\hat{\theta}(k)$ is the value $\theta \leq k$ that maximizes

$$l(k;\theta)=d'(k;\theta)C^{-1}(k;\theta)d(k;\theta) \quad (14)$$

and our decision rule is

$$l[k;\hat{\theta}(k)] \underset{H_0}{\overset{H_1}{\gtrless}} \epsilon \quad (15)$$

where ϵ is a threshold value chosen to provide a reasonable tradeoff between false and missed alarms (see discussion in Section V).

Note that the full implementation of the GLR detection-estimation scheme involves a growing bank of matched filters, i.e., we must compute $d(k;\theta)$ and $l(k;\theta)$ for $\theta=1,\cdots,k$. To avoid this problem, it is convenient to restrict attention to a "data window" of finite width. That is, at any time k, we restrict our optimization over θ to an interval of the form $k-M<\theta\leq k-N$. Note if $N=M-1$, $\theta=k-M+1$ there is no optimization over θ (the GLR consists of a single MF). If the window is sufficiently wide to insure detection and identification of all important failures, this approximation does not lead to serious difficulties. An algorithm based upon this finite window is developed in the Appendix. Means for selecting an appropriate window width are discussed in Section IV.

We now consider the case in which we hypothesize $\nu=\alpha f_i$ where α is an unknown scalar and $f_i\in\{f_1,\cdots,f_N\}$ is a given set of hypothesized "failure directions." In this case, the GLR detector takes the form

$$\hat{\alpha}(k)=\frac{b[k;\hat{\theta}(k),\hat{i}(k)]}{a[k;\hat{\theta}(k),\hat{i}(k)]} \quad (16)$$

where $\hat{\theta}(k)$ and $\hat{i}(k)$ are the quantities that maximize

$$l(k;\theta,i)=\frac{b^2(k;\theta,i)}{a(k;\theta,i)} \quad (17)$$

where

$$a(k;\theta,i)=f_i'C(k;\theta)f_i \quad (18)$$

$$b(k;\theta,i)=f_i'd(k;\theta). \quad (19)$$

The decision as to failure is made using the decision rule

$$l[k;\hat{\theta}(k),\hat{i}(k)] \underset{H_0}{\overset{H_1}{\gtrless}} \epsilon. \quad (20)$$

In general, $C(j;\theta)$, $G(j;\theta)$, $V(j)$, and the other relevant matrices described in the Appendix are time varying. However, if the system under consideration is time-invariant and if we use the optimal steady-state Kalman filter, we have that $C(j;\theta)=C(j-\theta;0)$, $G(j;\theta)=G(j-\theta;0)$, $V(j)=V(0)$, etc., which simplifies the necessary computation and storage. We note that in any case, the necessary quantities can be computed recursively (see the Appendix).

IV. Adaptive Filtering by Direct Estimate Incrementation and Covariance Incrementation

Once a jump has been detected by the GLR detector, we can use the MLE's $\hat{\theta}(k)$ and $\hat{\nu}(k)$ to directly increment our state estimate. We propose the estimate update equation

$$\hat{x}(k|k)_{\text{new}}=\hat{x}(k|k)_{\text{old}}+\{\Phi[k,\hat{\theta}(k)]-F[k;\hat{\theta}(k)]\}\hat{\nu}(k) \quad (21)$$

where Φ and F are defined in the Appendix. An adaptive filtering scheme based on this approximation, the finite data window method, and unconstrained ν is illustrated in Fig. 1. Once a state jump is detected, we increment the state estimate using (21). Note that the term $\Phi\hat{\nu}$ represents the contribution to $x(k)$ if a jump of $\hat{\nu}$ occurs at time $\hat{\theta}$, while $F\hat{\nu}$ represents the response of the Kalman–Bucy filter to the jump prior to its detection.

The proposed adaptive filtering scheme deserves some comment. We first note that it does not represent the optimal solution to the filtering problem for linear systems subject to jumps. Indeed, the feedback (21) uses only the MLE $\hat{\theta}(k)$ and totally neglects any information concerning the conditional distribution for θ. However, some simple reasoning yields the result that the estimate in (21) is precisely the optimal estimate for $x(k)$ given the measurements $z(j)$, $j=\hat{\theta}(k)$, $\hat{\theta}(k)+1,\cdots,k$ and the assumption that there is no *a priori* information (i.e., we assume that the *a priori* covariance for x at time $\hat{\theta}(k)$ is infinite).[1] This fact, combined with the simplicity of (21), provides the motivation for the use of this estimate incrementation method. We note that the suboptimality introduced by neglecting any uncertainty in $\hat{\theta}(k)$ may be of importance in some problems. In this case, one may wish to utilize more information about θ in the estimate update procedure. To this end, we remark that the $l(k;\theta)$ in (14) are likelihood ratios for various values of θ, and hence can be normalized to yield the conditional probability distribution for θ. We also note, however, that there is a critical tradeoff between filter performance and filter complexity, and this consideration has led us to propose computational simplifications such as the finite data window and the estimate update law (21).

We note that one could employ the estimate update procedure (21) without changing filter gains, thus avoiding on-line covariance calculations. However, in many cases the GLR estimate $\hat{\nu}(k)$ may be quite inaccurate, and the proposed scheme may lead to instabilities (we detect our inaccuracy in estimating ν as another jump and feed back another inaccurate estimate). Intuitively, we should increase our estimation error covariance to reflect the degradation in the quality of our estimate caused by the jump. By increasing the error covariance, the filter gain is increased and the filter can improve its response to the jump (i.e., it can compensate for inaccuracies in our estimates ν and θ). Given the interpretation of (21) as the optimal estimate based on $z[\hat{\theta}(k)],\cdots,z(k)$ with no *a priori* information, a reasonable method for covariance incrementation is to reset $P(k|k)$ to the error covariance for this estimate.

[1] This fact was pointed out to us by one of the reviewers.

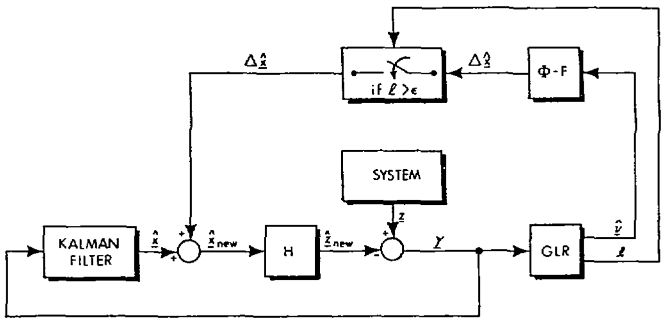

Fig. 1. Implementation of direct compensation technique.

Recalling that $C^{-1}[k;\hat{\theta}(k)]$ is the error covariance for $\hat{\nu}(k)$, we find that the appropriate covariance update equation is

$$P(k|k)_{\text{new}} = P(k|k)_{\text{old}} + \{\Phi[k,\hat{\theta}(k)] - F[k;\hat{\theta}(k)]\}C^{-1}[k;\hat{\theta}(k)] \cdot \{\Phi[k,\hat{\theta}(k)] - F[k;\hat{\theta}(k)]\}. \tag{22}$$

We close this section by making several more qualitative comments. We note that the incrementation procedure described by (21) and (22) requires the discarding of all filter information prior to $\hat{\theta}(k)$. This is a direct consequence of the assumption that ν has an infinite *a priori* covariance. In many problems, such as the inertial calibration and alignment problem discussed in [1], one has some *a priori* information about the possible range of values for ν. In addition, in problems in which great accuracy is required it does not seem appropriate to throw away all previous information. One possible method for avoiding such a problem is to use a finite initial covariance for ν. In addition, we have found that the use of gain incrementation (22) without estimate incrementation (21) also works well, as we essentially allow the filter to correct for the jump itself. This latter method can be interpreted as follows: prior to gain incrementation the filter bandwidth may be quite small and the filter responds very slowly to the jump; the use of gain incrementation then provides a mechanism for increasing the bandwidth and reducing the response time of the filter.

A final issue concerns the size M of the finite data window of the GLR. Noting the interpretation of $C^{-1}[k;\hat{\theta}(k)]$ as the error covariance for $\hat{\nu}(k)$, one might consider choosing a value of M such that this covariance is sufficiently small. On the other hand, in many cases one is much more concerned with quick *detection* of jumps rather than accurate estimation of the jump. In this case the crucial quantity that one wishes to control is *not* $C^{-1}[k;\hat{\theta}(k)]$ but rather the detection delay time $k-\hat{\theta}(k)$. This tradeoff between fast detection and accurate estimation is of obvious importance, and the relative importance of these issues in a particular application will dictate the choice of M.

Finally, we note that in the implementation of the GLR system of Fig. 1, we reinitialize the GLR detector (i.e., set the states of the MF's to zero) immediately following filter incrementation. This is done in order to avoid possible instabilities in which we detect and identify the same jump several times and "overcompensate." With this implementation, the GLR system can be used to detect several successive state jumps.

V. Detection Probability Calculation

Implementation of the GLR requires the choice of a decision threshold ϵ and a window length M. As mentioned earlier, these quantities are chosen by considering tradeoffs among detection delay time, the probability P_F of false alarm, and the probability $P_D(\nu,\theta)$ of correct detection of a jump of magnitude ν at time θ. These probabilities are given by [20]

$$P_F = \int_{\epsilon}^{\infty} p(l=L|H_0)\,dL \tag{23}$$

$$P_D(\nu,\theta) = \int_{\epsilon}^{\infty} p(l=L|H_1,\nu,\theta)\,dL. \tag{24}$$

Here $p(l=L|H_0)$ is the probability density of $l(k;\theta)$ conditioned on H_0, and $p(l=L|H_1,\nu,\theta)$ is its density conditioned on H_1 and particular assumed values of ν and θ. It is easy to show [22] that $P(l=L|H_0)$ is a Chi-squared density with n degrees of freedom, and $P(l=L|H_1,\nu,\theta)$ is a noncentral Chi-squared density [23], [24] with noncentrality parameter

$$\delta^2 = \nu^T C(k;\theta)\nu. \tag{25}$$

Values of P_F and P_D can be computed from tables in [23], as can the value ϵ for specified P_F or P_D. Finally, we note that in order to use these tables, we must specify values for θ and ν. A good guideline is to choose ν to be the minimum jump that must be detected for each failure direction. Also, several values of θ should be used if the system is time-varying.

VI. An Example

The application of the GLR to a simple tracking problem is illustrated in this section. The problem is to design a tracking filter which uses position measurements taken at 30 s intervals to track the motion of a vehicle along a straight line. The system dynamics are given by (1)–(4), where the only modeled force acting on the vehicle is a white Gaussian acceleration, $\Delta t = 30$ s, and

$$\Phi(k,k-1) = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix}, \quad \Gamma = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{26}$$

$$E[w(k)\;w(j)] = \delta_{kj}Q = (0.173\text{ ft/s})^2\delta_{kj}. \tag{27}$$

The measurement matrix and measurement noise are

$$H = [1\quad 0] \tag{28}$$

$$E[v(k)\;v(j)] = R\delta_{kj} = (600\text{ ft})^2\delta_{kj}. \tag{29}$$

The vehicle is subject to occasional step changes of unknown magnitude in either position or velocity. The tracking filter is a Kalman filter operating in steady state and requires 60–90 min to completely respond to such jumps. The magnitudes of the position and velocity jumps that are considered in this study are 10 times the steady-state rms estimation errors in the corresponding state variables.

The GLR system was implemented with the *detection* law

$$l(k;\theta) \underset{H_0}{\overset{H_1}{\gtrless}} 10.6 \tag{30}$$

and with $M=12$ and $N=6$, i.e., the optimization of $\hat{\theta}(k)$ was constrained to $k-12 < \hat{\theta}(k) \leqslant k-6$. Jump *identification* is made at the first time at which (30) is satisfied and

$$\hat{\theta}(k) = k - 11 \tag{31}$$

since waiting until the end of the detection window yields the most accurate estimate of $\hat{\nu}$ (see the preceding section). For the 10 σ jumps described above, (30) yields a P_F of 0.005 and a P_D greater than 0.9.

Fig. 2 presents a single Monte Carlo trial for the 10 σ (1320 ft) position jump. GLR detected the jump at 6.5 min (T_D) and identified it at 11 min (T_I). The constrained optimization of $\hat{\theta}(k)$ has the effect of implementing GLR as a finite memory filter. Thus, once θ is no longer in the detection window, i.e., $\theta < k-11$, *and* the tracking filter has responded sufficiently to the step change by itself, the detection law (30) becomes less likely to continue to select H_1. The tendency of $l(k;\theta)$ to decrease is evident in Fig. 2(b) for times greater than 16 min, although threshold crossings persist well past the 50 min mark. In 30 Monte Carlo trials with either the 10σ position or velocity jump, GLR correctly detected every jump.

Fig. 3 contains a summary of the jump estimation errors at T_D for a set of 10 Monte Carlo trials with the 1320 ft jump in position. In all but one trial, the error in $\hat{\theta}(k)$ was 30 s (one time step) or less. The errors in

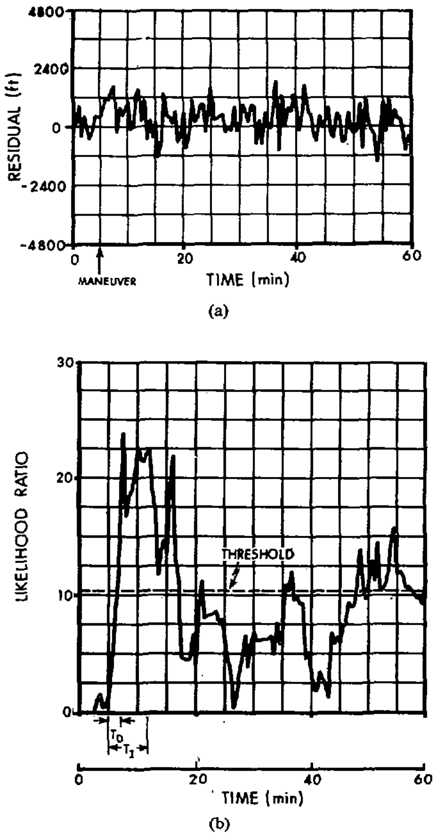

Fig. 2. (a) Filter residuals for a 1320 ft jump in position at 5 min. (b) Likelihood ratio for (a) using unconstrained GLR.

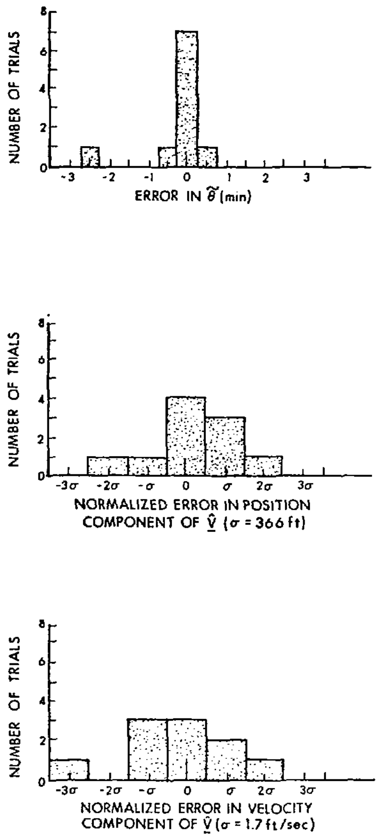

Fig. 3. Summary of jump estimation errors for Monte Carlo trials with a 1320 ft jump in position.

$\hat{\theta}(k)$ are normalized with respect to the σ's determined from $C^{-1}(k;k-11)$. The errors in $\hat{\nu}(k)$ appear to be approximately Gaussian with the predicted σ's.

Sets of Monte Carlo trials have been run in which the tracking filter estimates were compensated after a jump had been detected. The results are summarized in Table I (response time is the time required to reduce the estimation error to less than four times the steady-state standard deviation). The table indicates that direct compensation and gain compensation are both effective (compared to the response when no compensation is made), but for this example, gain compensation is superior.

VII. Conclusions

In this short paper we have developed an adaptive filtering technique for discrete-time linear stochastic systems subject to abrupt jumps in state variables. Our technique is potentially useful in the design of failure detection and compensation systems. The proposed estimation system consists of a Kalman–Bucy filter based on the "no-jump" hypothesis and a detection-compensation system based on generalized likelihood ratio (GLR) hypothesis testing. Once a jump is detected, we can adjust the filter in one of three ways: we can directly increment the state estimate using the parameter estimates provided by the GLR detector; we can increase the estimation error covariance using GLR data and thus can allow the filter to adjust itself to the jump; we can adjust both the estimate and the error covariance. A second-order example has been studied to indicate the dynamical characteristics of the GLR system. These results indicate the potential usefulness of the method, as extremely high correct detection rates and very small false alarm rates were obtained. We note that the structure of our system is appealing, especially in failure detection applications in which we do not wish to disrupt system performance until after a jump is detected. In addition, the GLR system can essentially be attached to the "end" of an existing filter that does not account for jumps.

The analysis of this short paper is devoted to the development and description of the algorithm and to a brief look at an analysis of its usefulness (i.e., the simulation results and the determination of expressions for the probabilities of false alarm and correct detection). Clearly further analysis and simulations are needed to assess the behavior of the overall system and to answer questions such as the stability of the filter, the choice of detection threshold, and the length of the detection window. It is our opinion that the GLR method will prove to be an extremely useful tool in the design of failure detection systems. In a later paper we will report on some extremely successful results for a problem involving the detection of gyro and accelerometer failures in an inertial navigation system.

Appendix
Computational Algorithm

Utilization of the GLR test (with $N=0$ for simplicity) requires that $l(k;\theta)$ in (14) be evaluated for all $k-M<\theta\leq k$. The necessary quantities for this calculation can be computed from the equations

$$C(k;\theta)=G'(k;\theta)V^{-1}(k)G(k;\theta)+C(k-1;\theta) \quad \text{(A1)}$$

$$d(k;\theta)=G'(k;\theta)V^{-1}(k)\gamma(k)+d(k-1;\theta) \quad \text{(A2)}$$

$$G(k;\theta)=H(k)[\Phi(k,\theta)-\Phi(k,k-1)F(k-1;\theta)] \quad \text{(A3)}$$

$$F(k;\theta)=K(k)G(k;\theta)+\Phi(k,k-1)F(k-1;\theta) \quad \text{(A4)}$$

where Φ is the state transition matrix for (1). Equations (A1)–(A4) are computed for $k-M<\theta<k$, and the only new additional equations are for a jump at the present time k.

$$C(k;k)=H'(k)V^{-1}(k)H(k) \quad \text{(A5)}$$

$$d(k;k)=H'(k)V^{-1}(k)\gamma(k) \quad \text{(A6)}$$

$$F(k;k)=K(k)H(k). \quad \text{(A7)}$$

If the system is time-invariant, (A1)–(A4) only need be utilized for the

TABLE I

SUMMARY OF FILTER RESPONSE TIMES FOR MONTE CARLO TRIALS WITH AND WITHOUT COMPENSATION

Series	No Compensation		Direct Compensation		Gain Compensation	
	Ave. Response Time (min)	Low-High (min)	Ave. Response Time (min)	Low-High (min)	Ave. Response Time (min)	Low-High (min)
Position (10σ)	54	44 - 71	34	10 - 57	22	12 - 39
Velocity (10σ)	49	39 - 61	34	10 - 81	16	4 - 40

first M measurements. For subsequent measurements the stored matrices from previous iterations can be used.

A last comment should be made concerning the existence of the inverses $C^{-1}(k;\theta)$. For an observable system there is some minimal integer N' such that $C^{-1}(k;\theta)$ does not exist for $\theta < k-N'$. This situation arises because the system is not completely observable from N' or fewer measurements. In this case, a reasonable approach is to choose some integer $N' \leqslant N < M$ and constrain the optimization to $k-M<\theta \leqslant k-N$. With this constraint, we need only store the corresponding values of the matrix functions defined in (A1)–(A4). In the special case in which N is equal to $M-1$, $\hat{\theta}(k)$ is $k-M+1$, and the optimization is eliminated. For unobservable systems, it is necessary to define pseudoinverses of $C^{-1}(k;\theta)$ which restrict the possible jump directions to some observable subspace of the state space.

ACKNOWLEDGMENT

The authors wish to express their appreciation to Prof. J. J. Deyst, Jr., of M.I.T. and to T. H. Kerr, B. S. Crawford, and E. M. Duiven of the Analytic Sciences Corporation for their many useful suggestions during the development of these results. The authors would also like to thank Major P. Pirtle for providing the motivation for this research effort. In addition, the authors are indebted to the reviewers for their many comments which have improved the presentation of the results of this short paper.

REFERENCES

[1] A. S. Willsky, J. J. Deyst, Jr., and B. S. Crawford, "Adaptive filtering and self-test methods for failure detection and compensation," in *Proc. of the 1974 Joint Automatic Control Conf.* June 19-21, 1974; also *J. Spacecraft Rockets*, vol. 12, pp. 434–437, July 1975.

[2] H. L. Jones, "*Failure detection in linear systems*," Ph.D. dissertation, Dep. Aeronaut. Astronaut., Massachusetts Inst. Technol., Cambridge, Sept. 1973.

[3] R. C. Montgomery and A. K. Caglayan, "A self-reorganizing digital flight control system for aircraft," presented at the AIAA 12th Aerospace Sciences Meeting, Jan. 30–Feb. 1, 1974.

[4] R. J. McAulay and E. Denlinger, "A decision-directed adaptive tracker," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-9, pp. 229–236, Mar. 1973.

[5] R. K. Mehra and J. Peschon "An innovations approach to fault detection and diagnosis in dynamic systems," *Automatica*, vol. 7, pp. 637–640, Sept. 1971.

[6] R. K. Mehra, "On the identification of variances and adaptive Kalman filtering," *IEEE Trans. Automat. Contr.*, vol. AC-15, pp. 175–184, Apr. 1970.

[7] T. J. Tarn and J. Zaborszky, "A practical nondiverging filter," *AIAA J.*, vol. 8, pp. 1127–1133, June 1970.

[8] A. H. Jazwinski, "Limited memory optimal filtering," *IEEE Trans. Automat. Contr.* (Short Papers), vol. AC-13, pp. 558–563, Oct. 1968.

[9] D. G. Lainiotis, "Joint detection, estimation and system identification," *Inform. Contr.*, vol. 19, pp. 75–92, Aug. 1971.

[10] D. G. Lainiotis and S. K. Park, "On joint detection, estimation and system identification: Discrete data case," *Int. J. Contr.*, vol. 17, no. 3, pp. 609–633, 1973.

[11] D. G. Lainiotis, "Optimal adaptive estimation: Structure and parameter adaptation," *IEEE Trans. Automat. Contr.*, vol. AC-16, pp. 160–170, Apr. 1971.

[12] T. Soeda and T. Yoshimura, "A practical filter for noisy dynamic systems with unknown time-varying parameters," in *1973 Joint Automatic Control Conf., Preprints*, pp. 415–421.

[13] M. Athans and D. Willner, "A practical scheme for adaptive aircraft flight control systems," presented at the Symp. Parameter Estimation Tech. and Appl. in Aircraft Flight Testing, NASA Flight Research Center, Edwards Air Force Base, CA, Apr. 24–25, 1973.

[14] A. H. Haddad and J. B. Cruz, Jr., "Non-linear filtering for systems with random models," presented at the Symp. Nonlinear Estimation Theory, San Diego, CA 1971.

[15] J. C. Deshpande, T. N. Upadhyay, and D. G. Lainiotis, "Adaptive control of linear stochastic systems," *Automatica*, vol. 9, pp. 107–115, 1973.

[16] P. Sanyal and C. N. Shen, "Bayes' decision rule for rapid detection and adaptive estimation scheme with space applications," *IEEE Trans. Automat. Contr.* (Short Papers), vol. AC-19, pp. 228–231, June 1974.

[17] D. D. Sworder and V. G. Robinson, "Feedback regulations for jump parameter systems with state and control dependent transition rates," *IEEE Trans. Automat. Contr.*, vol. AC-18, pp. 355–360, Aug. 1973.

[18] N. E. Nahi and B. M. Schaefer, "Decision-directed adaptive recursive estimators: Divergence prevention," *IEEE Trans. Automat. Contr.*, vol. AC-17, pp. 61–68, Feb. 1972.

[19] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. New York: Academic, 1970.

[20] H. L. Van Trees, *Detection, Estimation and Modulation Theory*. New York: Wiley, 1968, pt. 1.

[21] T. H. Kerr, "A two ellipsoid overlap test for realtime failure detection and isolation by confidence regions," presented at Pittsburgh Conf. Modeling and Simulation, Apr. 24–26, 1974.

[22] A. S. Willsky and H. L. Jones, "A generalized likelihood ratio approach to state estimation in linear systems subject to abrupt changes," in *Proc. 1974 IEEE Conf. Decision and Control*, Nov. 1974.

[23] D. B. Owen, *Handbook of Statistical Tables*. Reading, MA: Addison-Wesley, 1962.

[24] H. Scheffe, *The Analysis of Variance*. New York: Wiley, 1958.