Suboptimal Multi-Mode State Estimation and Mode Detection

Abdul Basit Memon∗ and Erik I. Verriest †
School of Electrical and Computer Engineering
Georgia Institute of Technology, Atlanta, Georgia 30332–0250
∗abmemon@gatech.edu, †erik.verriest@ece.gatech.edu

Abstract—A suboptimal state estimation algorithm is proposed for a discrete time stochastic multi-mode switched system, with a finite number of modes. The system state is transferred from one mode to the next. The parameters in the different modes are assumed to be known, but the mode-switching is random. The proposed filter is derived by extending the innovations approach to the problem of estimating the state of the system assuming that the switch between modes is purely random, but with known mode probabilities. As a special case we derive the filter for a linear system degraded by intermittent sensor failure and analyze its performance. This model is an alternative to the packet loss model, where the presence or absence of a packet is part of the observed data.

I. THE PROBLEM

Consider a finite set of linear time invariant discrete time systems. Let the set be indexed by $i$. Then each of these systems has dynamics described by:

$$x_{k+1}^{(i)} = F^{(i)}x_k^{(i)} + G^{(i)}w_k^{(i)}$$
$$y_k^{(i)} = H^{(i)}x_k^{(i)} + v_k^{(i)}.$$  

Let the noise sequences $\{w_k^{(i)}\}$ and $\{v_k^{(i)}\}$, in this model, be stationary white noises with zero mean and covariance matrices $Q^{(i)} \geq 0$ and $R^{(i)} > 0$ respectively and with cross covariance matrix $C^{(i)}$. The system and noise parameters for the $i$-th system will be denoted by $S_i$. Without loss of generality, we may assume that the dimensions are equal, i.e. $F^{(i)}, G^{(i)}, Q^{(i)} \in \mathbb{R}^{n \times n}, H^{(i)} \in \mathbb{R}^{n \times p}$, and $R^{(i)} \in \mathbb{R}^{p \times p}$. (A theory of multi-mode multi-dimensional systems ($M^3D$) is expounded in [19], [20]). In this paper we consider the case where the mode switching process is purely random, independent of the state sequence $\{x_k\}$, but with perhaps time varying occupation probabilities. Our compound systems constitute thus a specific class of stochastic hybrid or switched systems. A complete description of the state at time step $k$ of the compound system involves thus not only the state vector which is usually denoted by $x_k$, but also the mode $\theta_k$ which is active at time $k$.

Let $\pi_k = [\pi_1^k, \pi_2^k, \ldots, \pi_N^k]$ be the probability vector for these different modes, with $\pi_i$ being the probability that mode $i$ is ‘active’. An alternate description of these systems is presented here by a single $n$-th order system but with time varying parameters, whose values are determined by the ‘modal’ process $\{\theta_k\}$:

$$x_{k+1} = F(\theta_k)x_k + G(\theta_k)w_k$$
$$y_k = H(\theta_k)x_k + v_k.$$  

Also we have now $Q(\theta_k)$ and $R(\theta_k)$. Without loss of generality we state $Q^{(i)} = I$, since its variation can always be modeled in the $C^{(i)}$ matrices. In this model the modal state $\theta_k$ is a purely random sequence taking integer values. The probability that $\theta_k = i$ is $\pi_i^k$ and if $\theta_k = i$, then the system parameters are $F^{(i)} = F^{(i)}$ and so on. It is also important to note that at each time, only one output vector is available corresponding to the mode at that time ($y_k = y^{(\theta_k)}$).

The objective here is to obtain a good causal estimate of the state process $\{x_k\}$ from the past of the observed sequence $\{y_k\}$ and perhaps some prior information - the occupation probabilities in our case. Although this problem can be formulated in the exact least squares sense, with the conditional expectation as solution, finding this optimal solution is prohibitively complex due to the Bayesian explosion. (See [7], [13], [10], [9], [6] and more recently [4], [3]). Therefore, the objective of our work is to obtain a simple suboptimal scheme for the estimation of state sequences of multi-mode systems with random switching. Towards that end, we will be deriving the linear least squares filter for this problem ([18], [11]).

A. Relevance of this Problem and Related Work

Multi-mode systems with random switching appear in many areas such as networked control systems, where the jumping characterizes packet loss or delay ([2]). These problems are also of interest in sensor networks. Sensing may be performed at locations that are geographically remote from the central processor (See [21], [15], [8]). Hence communication, and with it communication constraints become an integral part of the estimation problem. Estimation in a network with packet losses is a problem of this scenario: Either normal operating conditions for packet delivery prevail, or there is a failure in packet transmission. In [8], a Markovian model was assumed for switching between these two modes. Their emphasis is however not on the optimal filtering but rather an analysis of the stability. Sufficient conditions are given in terms of the failure and recovery rates (for a first order system, the condition also turns out to be necessary). See also [10], [14]. Sinopoli et al. [2004] consider and analyze the case of intermittent observations and derive a threshold value for the arrival rate, below which divergence occurs. The problem of intermittent observations is also discussed in detail in this paper. While Sinopoli et al. ([15], [16]) assume complete knowledge of the presence or
absence of the signal in the observations, we emphasize that this information is a hidden variable in our model and this sets our work apart from the theory of Sinopoli et al. (Just before submitting this work, reference [12] appeared.)

Switching Kalman filters are further employed in voice activity detection schemes ([5]), in navigation and tracking ([11]), and in vehicle tracking with vision ([17]). With a view towards improved prosthetics, [22] models motor cortical activity using a switching Kalman filter. Typical estimation schemes employed in these cases involve Kalman filter banks such as the Interacting Multiple Models (IMM) algorithm. While the problem at hand could be considered as an adaptive filtering problem, the structure (typically slow adaptation) may not be suitable for fast intermittency or switching in the model set. In addition, adaptive filtering typically deals with non discretely varying system parameters.

The paper is organized as follows. In section II, we derive the Multi Mode Filter for purely random mode transitions and discuss its computational complexity. In section III, we consider the special case of intermittent signal observations in noise in great detail and analyze the performance of our filter and compare it to other schemes.

II. THE MULTI MODE FILTER

We shall be interested in linear least squares filters and so the estimates we shall obtain are not the optimal least squares estimates, which are conditional expectations, given the data. However, the computational advantages for the suboptimal linear least squares solutions are great (as well known for the classical Kalman filter for linear stochastic system with non gaussian initial conditions). Of the two typical approaches to filtering - the innovations approach and the change of measure method we will use the innovations approach. The approach requires the Hilbert space framework for the stochastic processes involved, and the linear least squares estimate given the data is known to be a linear projection onto the subspace spanned by the data. For this reason, let us first classify the linear information structures.

Let \((\Omega, B, P)\) be the probability space on which all random variables involved in the state space model are defined: i.e., in particular \(x_0 \in B\), and for all \(k, \theta_k \in B, w_k \in B\) and \(v_k \in B\). The following linear spaces play an important role in the linear estimation problem:

\[
\mathcal{L}^0_k = \text{span}\{x_0, \theta_0, \ldots, \theta_k, w_0, \ldots, w_k, v_0, \ldots, v_k\} \quad (5)
\]

\[
\mathcal{L}_k = \text{span}\{\theta_0, \ldots, \theta_k, w_0, \ldots, w_k, v_0, \ldots, v_k\} \quad (6)
\]

\[
\mathcal{H}_k = \text{span}\{y_0, \ldots, y_k\}. \quad (7)
\]

The last space is generated by the observations. Define the innovations, \(\epsilon_k = y_k - \hat{y}_{k|k-1}\), where \(\hat{y}_{k|k-1}\) is the projection of the measurement \(y_k\) onto the subspace \(\mathcal{H}_{k-1}\). Its computation is simplified by the smoothing formula (9)

\[
\mathcal{P}^{\mathcal{H}_{k-1}|y_k} = \mathcal{P}^{\mathcal{H}_{k-1}|\mathcal{L}_{k-1}} [H(\theta_k)x_k + v_k] \]

\[
= \mathcal{P}^{\mathcal{H}_{k-1}} \left[ \sum \pi_i H^{(i)} x_k \right] \quad (13)
\]

If \(\chi_i(\theta)\) is the indicator function for mode \(i\), then

\[
\mathcal{P}^{\mathcal{H}_{k-1}|H(\theta_k)} = \sum \chi_i(\theta_k) H^{(i)} = \sum \pi_i H^{(i)} \quad (14)
\]

1 A correction to the printed version was given in the presentation.
Thus the innovations are given by

\[
e_k = y_k - \sum \pi_i H^{(i)} \hat{x}_{k|k-1}
\]

\[
= H(\theta_k) \hat{x}_{k|k-1} + v_k + \left[ H(\theta_k) - \sum \pi_i H^{(i)} \right] \bar{x}_{k|k-1}
\]

where \( \bar{x}_{k|k-1} = x_k - \hat{x}_{k|k-1} \).

The innovations covariance is obtained by ‘squaring up’ and using (11)

\[
\dot{R}_k' = \mathbb{E} \epsilon_k' \epsilon_k^T
= \dot{H} P_{k|k-1} \dot{H}' + \sum \pi_i H^{(i)} \Pi_k \left[ H^{(i)} - \dot{H}' \right]' + \dot{R}
\]

The Kalman Gain is

\[
\mathbb{E} \epsilon_k' \epsilon_k^T = P_{k|k-1} \dot{H}'
\]

\[
\dot{\Pi}_k = \Sigma_{k|k-1} + P_{k|k-1} \dot{H}' R_k^- \epsilon (y_k - \dot{\Pi} \bar{x}_{k|k-1})
\]

\[
\text{The Kalman Gain is}
\]

\[
\mathbb{E} \epsilon_k' \epsilon_k^T = P_{k|k-1} \dot{H}'
\]

\[
\text{So,}
\]

\[
\dot{\hat{x}}_{k|k} = \hat{x}_{k|k-1} + P_{k|k-1} \dot{H} R_k^- \epsilon (y_k - \dot{\Pi} \bar{x}_{k|k-1})
\]

\[
\text{iv) Predicted Estimate from the Signal Model}
\]

\[
\mathcal{P} \mathcal{H}_k x_{k+1} = \mathcal{P} \mathcal{H}_k [F(\theta_k) x_k + G(\theta_k) w_k]
\]

Consider first,

\[
\mathcal{P} \mathcal{H}_k F(\theta) x_k = \mathcal{P} \mathcal{H}_k F(\theta) x_k + \mathcal{P} \mathcal{span}(v_k) F(\theta) x_k
\]

\[
= \left( \sum \pi_i F^{(i)} \right) \hat{x}_{k|k-1}
+ \mathbb{E} [F(\theta_k) \epsilon_k' \epsilon_k^T (y_k - \dot{\Pi} \bar{x}_{k|k-1})]
\]

\[
= \mathcal{F} \hat{x}_{k|k-1} + \left[ \sum \pi_i F^{(i)} P_{k|k-1} H^{(i)} \right] (y_k - \dot{\Pi} \bar{x}_{k|k-1})
\]

\[
= \mathcal{F} \hat{x}_{k|k-1} + \left[ \mathcal{F} P_{k|k-1} \dot{H}' - \mathcal{F} \Pi \dot{H}' \right] R_k^- \epsilon_k
+ \sum \pi_i F^{(i)} \Pi_k H^{(i)} R_k^- \epsilon_k
\]

where we also defined

\[
\mathcal{F} = \sum \pi_i F^{(i)}
\]

as the average dynamics. Consider next,

\[
\mathcal{P} \mathcal{H}_k G(\theta) w_k = \mathcal{P} \mathcal{span}(v_k) G(\theta) w_k
\]

\[
= \left[ \sum \pi_i G^{(i)} C^{(i)} \right] R_k^- \epsilon_k
\]

\[
\text{v) \ Finally, we obtain a recursion for the error covariance.}
\]

\[
\Sigma_{k+1} = \mathbb{E} [F(\theta_k) x_k + G(\theta_k) w_k] [F(\theta_k) x_k + G(\theta_k) w_k]' - \mathcal{F} \Sigma_{k|k-1} \mathcal{F} + \sum \pi_i G^{(i)} C^{(i)}
\]

\[
= \sum \pi_i \left( F^{(i)} \Pi_k H^{(i)} + G^{(i)} C^{(i)} \right) R_k^- \epsilon_k
\]

\[
\text{Likewise the covariance of the estimate follows by ‘squaring up’ the predictor equation (26). We find that } \Sigma_{k+1|k} \text{ equals}
\]

\[
\Sigma_{k+1|k} = \mathcal{F} \Sigma_{k|k-1} \mathcal{F} + \sum \pi_i \left( F^{(i)} \Pi_k H^{(i)} + G^{(i)} C^{(i)} \right) R_k^- \epsilon_k
\]

\[
\text{The decomposed equations give}
\]

\[
\Sigma_{k+1|k} = \sum \pi_i \Sigma_{k|k-1}
\]

where

\[
\Sigma_{k+1|k} = \mathcal{F} \Sigma_{k|k-1} \mathcal{F} + \sum \pi_i \left( F^{(i)} \Pi_k H^{(i)} + G^{(i)} C^{(i)} \right) R_k^- \epsilon_k
\]

\[
\Sigma_{k+1|k} = \sum \pi_i \left( F^{(i)} \Pi_k H^{(i)} + G^{(i)} C^{(i)} \right) R_k^- \epsilon_k
\]

\[
\text{Subtracting both equations yields the recursion for the error covariance:}
\]

\[
P_{k+1|k} = \Pi_{k+1|k} - \Sigma_{k+1|k}
\]

\[
\Pi_{k+1|k} = \Sigma_{k+1|k} - \Sigma_{k-1|k}
\]

Unfortunately, no decoupled recursions (in the form of Riccati equations) exists for the } P_{k+1|k}.
We summarize:

**Theorem 2.1:** The linear least squares filter for a multi-mode system (MMF) with purely random mode switching for the system given by (3) and (4) is given by (recall: $Q = I$)

$$\hat{x}_{k+1|k} = F\hat{x}_{k|k-1} + K_k (y_k - \Pi\hat{x}_{k|k-1})$$

$$K_k = (FP_k|_{k-1} \Pi + GG' + F\Pi_k\Pi' - F\Pi_k\Pi')^{-1}R_k^{-\epsilon}$$

$$\Pi_{k+1} = F\Pi_kF' + GG'$$

$$P_{k+1|k} = FP_{k|k-1}F' + GG' - K_k\Pi_kR_k^{-\epsilon}K_k' + F\Pi_kF' - F\Pi_k\Pi'$$

The filtered estimates $\hat{x}_{k|k}$ follow from

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + P_{k|k-1}^{\epsilon} \Pi_{k}^{-\epsilon} (y_k - \Pi\hat{x}_{k|k-1})$$

The ‘overline’ indicates expectation with respect to $\theta$, i.e., $\bar{X} = \sum_{i} \pi_i X^{(i)}$. It should be emphasized that this multi-mode filter is not the Kalman filter for the averaged system.

### A. Discussion

In this section, we will discuss the computational complexity of the multi-mode filter (MMF) when compared with algorithms using filter banks such as the IMM algorithm, typically employed to estimate the state of a multi-mode system. We compared the number of scalar multiplication and addition operations required by full implementations of the MMF with a Kalman filter bank. We found out that the number of operations required by MMF is dominated by $n^3$, whereas the number of operations required by a filter bank is dominated by $n^2 \cdot p^3$. Here $n$ is the dimension of the state space, $p$ is dimension of the output space and $q$ is the number of modes in the system. The MMF is found to be computationally simpler for large number of modes or large number of outputs. It is also worth noting that if the steady state covariance of the system is used in the MMF equations to find an approximate filter, the MMF is always computationally simpler than a filter bank. A possible scheme that can be employed is to begin with a filter bank algorithm such as the IMM and then switch to the MMF after convergence to steady state. The MMF can also be used adaptively in the case of fixed but unknown occupation probabilities. A detection scheme can be created based on the residuals and the probabilities can be estimated based on the mode detected.

### III. Intermittent Signal Observations in Noise

This is a special case of the general problem treated above, and relevant in the sensor failure problem. Consider the system,

$$x_{k+1} = Fx_k + Gw_k$$

$$y_k = \theta_k Hx_k + v_k.$$  \hspace{2cm} (35)

The $\{\theta_k\}$ is a scalar Bernoulli process, with value space $\{0, 1\}$. Let $\pi_1(k) = \Pr \{ \theta_k = 1 \}$. Making this process nonstationary allows for a model incorporating faulty behavior of the sensor as function of time. In this case there are two modes for the system. In mode 1, the parameters are $S_1 = \{ F, G, H, Q = I, R > 0 \}$, whereas the mode 0, the sensor failure mode, is characterized by $S_0 = \{ F, G, 0, Q = I, R \}$. The probabilistic description of the modes is given by $\pi(k) = [\pi_1(k), \pi_0(k)] = [\lambda_k, 1 - \lambda_k]$. Additionally, the two noise sequences $w_k^{(i)}$ and $v_k^{(i)}$ are uncorrelated for all time i.e., $C^{(i)} = 0$.

It follows from Theorem 2.1, with the averaged parameters $\bar{F} = F$, $\bar{H} = \lambda_k H$, $GG' = GG'$, $\bar{R} = R$, that the MMF is (with $Pr\{sensor\ failure\} = 1 - \lambda_k$

$$\hat{x}_{k+1|k} = F\hat{x}_{k|k-1} + K_k (y_k - \lambda_k H\hat{x}_{k|k-1})$$

$$K_k = \lambda_k FP_{k|k-1} H' \Pi_k^{-\epsilon}$$

$$\Pi_{k+1} = F\Pi_kF' + GG'$$

$$P_{k+1|k} = FP_{k|k-1}F' + GG' - K_k\Pi_kR_k^{-\epsilon}K_k'$$

$$R_k = \lambda_k^2 H P_{k|k-1} H' + R + \lambda_k (1 - \lambda_k) H \Pi_k H'$$

### A. Stability

The following two theorems characterize the stability of this filter for the case of a stationary Bernoulli process.

**Theorem 3.1:** For the intermittent signal observations in the noise case, if the system is asymptotically stable then the multi-mode filter is stable i.e. the error covariance matrix $P_{k+1|k}$, given by the following recursion

$$P_{k+1} = FP_{k|k-1}F' + GG' - \lambda^2 FP_{k|k-1}H'$$

$$[\lambda^2 H P_k H' + R + \lambda (1 - \lambda) H \Pi_k H']^{-1} H P_k F'$$

converges to a steady state $P$ which satisfies an algebraic Riccati equation,

$$P = FPF' + GG' - FPH'[HP'H + R_{eq}]^{-1} HPF'$$

where,

$$R_{eq} = \frac{R + \lambda (1 - \lambda) H \Pi H'}{\lambda^2}$$

and II satisfies an algebraic Lyapunov equation.

$$\Pi = F\Pi F' + GG'$$

**Theorem 3.2:** For the intermittent signal observations in noise case, if the system is unstable and the pair $(F, H)$ is observable then the multi-mode filter is unstable i.e. the error covariance matrix does not converge to a steady state.

**Proof:** The Riccati recursion for the prediction error covariance matrix is given below where the notation $P_k$ has been used instead of $P_{k|k-1}$ to save space.

$$P_{k+1} = FP_{k|k-1}F' + GG' - \lambda^2 FP_{k|k-1}H'$$

$$[\lambda^2 H P_k H' + R + \lambda (1 - \lambda) H \Pi_k H']^{-1} H P_k F'$$

$$= GG' + F \{ P_k - P_k \lambda H' \}$$

$$[\lambda H P_k H' + R + \lambda (1 - \lambda) H \Pi_k H']^{-1} \lambda H P_k \} F'$$

$$= F \{ P_k^{-1} + \lambda H' [R + \lambda (1 - \lambda) H \Pi_k H'^{-1} \lambda H P_k \}^{-1} F' + GG'$$
where the last form is obtained by using Woodbury’s lemma on the expression enclosed in the curly brackets. If $F$ is unstable and $(F, H)$ is observable then the term $H\Pi_kH'$ is unbounded. This results in the term $[R+\lambda(1-\lambda)H\Pi_kH']^{-1}$ approaching 0 as $k$ approaches infinity. Thus, in the long run the Riccati equation behaves as the Lyapunov equation and $P_k$ is unbounded for the unstable system.

**B. Bounds on $P$**

An upper bound and a lower bound for the error covariance matrix $P_{k+1|k}$, hereafter denoted by $P_{k+1}$, can be easily found using Riccati equation comparison theorems due to Wimmer and Pavon ([23]).

**Theorem 3.3:** The error covariance matrix $P_k$ is bounded below for all time by the following Riccati recursion,

$$L_{k+1} = FL_kF' + GG' - FL_kH'(HL_kH' + R)^{-1}HL_kF'$$

provided $L_0 \leq P_0$.

**Proof:** It follows directly from Wimmer and Pavon’s comparison theorem if the following required condition is satisfied for all time,

$$
\begin{bmatrix}
GG' \\
F' - H'\tilde{R}_k^{-1}H
\end{bmatrix} \geq 
\begin{bmatrix}
GG' & F' \\
F' - H'R^{-1}H
\end{bmatrix}
$$

where $\tilde{R}_k = \lambda_k^{-2}[R+\lambda_k(1-\lambda_k)H\Pi_kH']$. Since the matrices $R$ and $\tilde{R}_k$ are symmetric, this condition is equivalent to $\tilde{R}_k \geq R$. Given that $\lambda_k \leq 1$ and $\Pi_k \geq \beta$ for all $k$, the inequality $\tilde{R}_k \geq R$ is true for all time and thus the bound holds.

**Theorem 3.4:** The error covariance matrix $P_k$ is bounded above for all time by the following Riccati recursion,

$$U_{k+1} = FU_kF' + GG' - FU_kH'(HUL_kH' + R_{eq})^{-1}HU_kF'$$

where $R_{eq} = \lambda_k^{-2}[R+\lambda_k(1-\lambda_k)H\Pi_kH']$ and $\Pi = F\Pi F'$ and $\Pi = GG'$ and provided $U_0 \geq P_0$.

**Proof:** Again this follows directly from Wimmer and Pavon’s comparison theorem if the following required condition is satisfied for all time,

$$
\begin{bmatrix}
GG' \\
F' - H'\tilde{R}_k^{-1}H
\end{bmatrix} \geq 
\begin{bmatrix}
GG' & F' \\
F' - H'\tilde{R}_k^{-1}H
\end{bmatrix}
$$

where $\tilde{R}_k = \lambda_k^{-2}[R+\lambda_k(1-\lambda_k)H\Pi_kH']$. Since the matrices $R_{eq}$ and $\tilde{R}_k$ are symmetric, the above condition is equivalent to $R_{eq} \geq \tilde{R}_k$. And this inequality is true for all $k$, since $\Pi_k \leq \beta$ for all $k$. Therefore, the upper bound holds.

**C. Numerical Example**

The performance of the MMF, for the intermittent signal in observations case with stationary Bernoulli process, is compared numerically to the IMM algorithm, the averaged filter, the approximate MMF and the exact knowledge case. The averaged filter is a Kalman filter for the averaged case, specifically $\bar{F} = \lambda H$. The exact knowledge case has perfect knowledge of the mode of operation at all times and employs the respective Kalman filter. The approximate MMF uses the system’s steady state covariance (II) to make the computations simpler. The performance metric used is the ratio of the euclidean norm of the error between the actual state and the estimated one to the norm of the error for the exact knowledge case. For each experiment, 1000 statistical trials are performed and the figures depict the computed ratios using statistical mean of the norm. Figure 1 compares the performance of the different filters for the system with the parameters $Q = 1$, $R = 25$, Sensor failure probability $= 0.3$ and

$$
F = \begin{bmatrix}
0.3 & 0 & 0 \\
0 & 0.7 & 0.7 \\
0 & -0.7 & 0.7
\end{bmatrix},
G = \begin{bmatrix}
5 \\
5
\end{bmatrix},
H = [1, 0, 0]
$$

The error is highest for the averaged filter followed by the MMF and then the IMM algorithm. It is also worth noting that the difference in the error norms is small and the approximate MMF, which is computationally simpler, performs as well as the full MMF.

![Fig. 1: Performance comparison between different filters for a typical stable case](image1)

Figure 2 illustrates the behavior of the filters when the system is close to instability i.e. one of the eigen values of the system is moved to 0.99 while keeping all the other parameters the same as the previous case. The trend observed in the previous case continues but now the difference in the scale of the norm between the MMF and averaged filter is much larger. The difference between the results of MMF and the IMM algorithm is slightly higher as well.

![Fig. 2: Performance comparison between different filters close to instability](image2)
D. Mode detection for adaptive filtering

If the occupation probability $\lambda_k$ is unknown then a simple mode detection scheme can be designed which, coupled with the MMF, estimates the state of the system adaptively. This mode detection scheme is outlined as follows. The residuals for both possible cases - signal in the observation and just noise in the observation, are computed at each time step.

$$\epsilon_k^1 = y_k - H\hat{x}_{k|k-1} \quad \text{and} \quad \epsilon_k^0 = y_k$$  \hspace{1cm} (37)

The Euclidean norms of these residuals are then computed and the most likely mode is determined as the one corresponding to the smaller of the norms of the two residuals.

$$\theta_k = \arg \min \{ \| \epsilon_k^1 \|, \| \epsilon_k^0 \| \}$$

The probability $\lambda_k$ can then be estimated by taking a moving average or a weighted moving average of the mode variable $\theta_k$. Finally, the estimated occupation probability $\hat{\lambda}_k$ is fed into the MMF to estimate the state of the system.

This scheme is motivated as follows. Firstly, we can rewrite (37) as,

$$\epsilon_k^0 = \theta_k H\hat{x}_{k|k-1} + v_k + \theta_k H\hat{x}_{k|k-1}$$  \hspace{1cm} (38)

$$\epsilon_k^1 = (\theta_k - 1) H\hat{x}_{k|k-1} + v_k + \theta_k H\hat{x}_{k|k-1}$$  \hspace{1cm} (39)

We define a new linear space $G_k = H_{k-1} \cup \{ \theta_k \}$. Note that,$$
\mathcal{P}^{G_k}_1 = \theta_k H\hat{x}_{k|k-1} \quad \text{and} \quad \mathcal{P}^{G_k}_0 = (\theta_k - 1) H\hat{x}_{k|k-1}
$$
and the randomness in both the residuals with respect to the space $G_k$ is the same i.e. $(v_k + \theta_k H\hat{x}_{k|k-1})$.

**Theorem 3.5:** $Pr[\mathcal{P}^{G_k}_1 \| \epsilon_k^0 \|^2 \geq \mathcal{P}^{G_k}_0 \| \epsilon_k^1 \|^2] = \lambda$.

**Proof:** Using (38) and (39), the left hand side of the above theorem can be easily worked out to be,

$$Pr[\mathcal{P}^{G_k}_1 \| \epsilon_k^0 \|^2 \geq \mathcal{P}^{G_k}_0 \| \epsilon_k^1 \|^2] = Pr[(1 - \theta_k) H\Sigma_{k|k-1} H^T \leq 0] = Pr[(1 - \theta_k) \leq 0] = Pr[\theta_k \geq \frac{1}{2}] = Pr[\theta_k = 1] = \lambda.$$

The reader might note that the aforementioned scheme uses just the Euclidean norms of the residuals (38) and (39) to find the estimate $\hat{\lambda}_k$. We conjecture that this scheme can be justified if $\lambda_k$ varies slowly and the averaging is over many samples.

IV. CONCLUSIONS

We extended the innovations approach to derive the linear least squares filter for a purely random multi mode or hybrid system. The methodology was applied to obtain the linear filters for systems with intermittently failing sensors (producing signal plus noise, or just noise). This special case was analyzed in detail and its performance compared favorably to other schemes. A detection scheme was also suggested to obtain estimates for the occupation probability $\lambda$ when the latter is unknown.

REFERENCES


