2010 IEEE 26-th Convention of Electrical and Electronics Engineers in Israel

A New Formulation of Fault-Tolerant Estimation Problems and Some Solutions

Daniel Sigalov
Interdepartmental Program for Applied Mathematics
Technion – Israel Institute of Technology

Yaakov Oshman
Department of Aerospace Engineering
Technion – Israel Institute of Technology

***Abstract*—The problem of fault tolerant state estimation is considered. We propose a unified, general formulation of the problem in which two different types of faults affect the system's output simultaneously. This problem statement generalizes previously reported formulations that may be obtained as special cases. Three families of state estimation methods for fault-prone systems are presented, generalizing several classical estimation algorithms. These families include: one-step near-optimal filters, which are closely related to the GPB filter family, IMM-based filters, and linear optimal estimators.**

I. Introduction

Fault tolerant state estimation is vital in many applications and especially in navigation systems. Playing a key role in operations such as aircraft steering and landing, spacecraft pointing and rendezvous missions, and missile guidance, navigation systems require reliable mechanisms for fault detection and isolation (FDI).

Although the most general formulation of fault-prone systems assumes that faults can alter both the dynamics equation as well as the measurement equation of the state-space system representation, an important subclass of problems assumes states that are unaffected by faults, such that fault indicators affect directly only the measurement equation. Typical examples are navigation systems using GPS signals that are prone to jamming and spoofing, and inertial sensors (accelerometers and rate gyros) that are, frequently, of low grade.

Two main types of faults characterize fault-prone systems. The first one is *multiplicative* faults, usually referred to as interruption fault indicators. These faults are modeled as Bernoulli random variables multiplying the state vector in the measurement equation. Thus, depending on the actual value of the fault indicator, the state may either be observed through the usual measurement channel, or go undetected such that the acquired signal carries no valuable information. The second type of faults is *additive* faults, also known as measurement biases. These are modeled as random additive disturbances, affecting, e.g., the nominal measurement noise variance.

Many estimation techniques consider systems affected by only one of the fault types, proposing various state estimation algorithms for such systems. One of the earliest contributions considering multiplicative faults is the work of Nahi [1], that examined the case, where the fault indicators constitute an i.i.d. Bernoulli process, and derived the linear optimal state estimator. Sawaragi et al. [2] considered a similar problem for correlated fault indicators at subsequent times. Hadidi and Shwartz [3] generalized the problem and considered general Bernoulli sequences, including a special case of a Markov sequence. Costa [4] considered a more general problem at the expense of state augmentation. For contributions involving the second type of faults the reader is referred to [5] and references therein.

Although different in nature, both types of fault processes may be interrelated, such that a fault event may result in an increased bias, misdetected measurement, or both. Thus, a general formulation requires a unified treatment of both kinds of faults affecting the system simultaneously. In [6] both types of faults were considered simultaneously, but these were governed by the same stochastic process, such that at a given time either both faults occurred or none.

We propose a general formulation of linear, fault-prone dynamical systems, that are subject to simultaneously-acting multiplicative and additive measurement faults. The formulations of [1], [3], and [6], as well as several additional problems, follow as special cases of our general formulation. We propose three families of suboptimal filters for state estimation in the above problems. These families are one-step optimal filters, that are closely related to the generalized pseudo-Bayesian (GPB) filter family [7], IMM-based filters [8], and linear optimal filters. For the latter case, the proposed algorithms relax several restrictive assumptions made in [3].

The remainder of this paper is organized as follows. In Section II we formulate the general fault-prone dynamical system and derive several special cases to be considered. In Sections III and IV, respectively, we derive one-step near-optimal and IMM-based filters for the general problem, and present closed-form algorithms for the special cases defined in Section II. In Section V the linear optimal filter is presented. Several numerical examples are presented in Section VI. Concluding remarks are made in Section VII.

II. Problem Description

Consider the following hybrid dynamical system

$$\mathbf{x}_{k+1} = A\mathbf{x}_k + \mathbf{w}_k \tag{1}$$

$$\mathbf{z}_k = \gamma_k H\mathbf{x}_k + \mathbf{v}_k(\delta_k). \tag{2}$$

$\{\mathbf{w}_k\}$ is a zero mean white Gaussian sequence with known covariance $Q_k \equiv Q$, $\{\gamma_k\}$ is a discrete-time Bernoulli random process, and $\{\delta_k\}$ is a sequence of positive binary random variables. Given $\{\delta_k\}$, $\{\mathbf{v}_k\}$ is a zero-mean white Gaussian sequence, such that $\operatorname{cov}(\mathbf{v}_k \mid \delta_k) = R\delta_k$. For simplicity we assume that δ_k may assume two values, 1 and $C > 1$, corresponding to the nominal and faulty situations, respectively. The process noise sequence is assumed to be independent of $\{\gamma_k\}$, $\{\delta_k\}$, and $\{\mathbf{v}_k\}$. All the sequences under consideration are independent of the initial state $\mathbf{x}_0$ which is assumed to be Gaussian with zero mean and known covariance P_0.

The described system is subject to two types of faults. The first type, associated with the sequence $\{\gamma_k\}$, is the multiplicative fault sequence, and the second type, associated with $\{\delta_k\}$, is the additive fault sequence. These two types of fault sequences correspond to the following undesirable effects. Under fault occurrence, the signal of interest, $\mathbf{x}_k$, may not be present in the acquired signal $\mathbf{z}_k$ at all, or may be corrupted by a stronger than nominal measurement noise. In the former case, the signal would carry a non-informative data of nominal or increased intensity.

978-1-4244-8682-3/10/$26.00

Whenever a multiplicative fault occurs, $\gamma_k = 0$ and the measured signal carries only noise. One would expect the filter to ignore the non-informative measurements and utilize only those for which $\gamma_k = 1$. Whenever an additive fault occurs, a faulty measurement results in a higher variance of the noise variable v_k. In this case a good filter is expected to use all measurements, assigning smaller weight to those carrying stronger noise.

Formulation (1)–(2) induces a family of problems, four of which we consider here:

Problem 1. $\{\gamma_k\}$, $\{\delta_k\}$ are mutually independent i.i.d. sequences.

Problem 2. $\{\gamma_k\}$, $\{\delta_k\}$ constitute a vector-valued i.i.d. sequence $\{\gamma_k, \delta_k\}$ with known joint distribution at time k, $\Pr\{\gamma_k = \gamma, \delta_k = \delta\}$.

Problem 3. $\{\gamma_k\}$, $\{\delta_k\}$ are mutually independent homogeneous Markov sequences with known transition laws.

Problem 4. $\{\gamma_k\}$, $\{\delta_k\}$ constitute a vector-valued homogeneous Markov sequence $\{\gamma_k, \delta_k\}$ with known joint transition kernel, $\Pr\{\gamma_k = \gamma, \delta_k = \delta \mid \gamma_{k-1} = \gamma_1, \delta_{k-1} = \delta_1\}$.

These problems may be viewed as listed in order of increasing complexity. In fact, Problem 1 is a special case of Problems 2–4. Problems 1–3 may be viewed as special cases of Problem 4. Note, however, that Problems 2 and 3 do not generalize each other.

As is well known, the optimal estimator $\hat{\mathbf{x}}_k$ using the available data $Z_k \triangleq \{\mathbf{z}_0, ..., \mathbf{z}_k\}$ is given by the conditional expectation $\mathbb{E}[\mathbf{x}_k|Z_k]$. The computation of this expectation requires an exponentially growing memory and is, therefore, impractical for all of the above problems. Consequently, we consider suboptimal approaches.

First, we consider recursive estimators that compute the estimate at time k, using the previously obtained estimate $\hat{\mathbf{x}}_{k-1}$ and the new measurement $\mathbf{z}_k$. Each such computation is performed by conditioning on the current fault indicators, which results in an optimal or near-optimal update with respect to this conditioning. The overall estimate, with respect to the entire measurement sequence, is suboptimal, due to the performed history merging. These estimators will be called in the sequel one-step near-optimal filters. The technique resembles GPB estimation, in that history sequences are merged and represented by a single estimate. In the second approach, we utilize the IMM method by using interaction of several primitive Kalman filters matched to different models. Finally, we derive linear optimal recursive estimators for the considered problems.

III. One Step Near-Optimal Filters

Consider the MMSE-optimal estimator of $\mathbf{x}_k$ using all the available measurements up to and including time k. It is given by the conditional expectation of the state given the measurements $Z_k \triangleq \{\mathbf{z}_0, ..., \mathbf{z}_k\}$. We expand the conditional expectation by conditioning on the latest joint fault indicator $\{\gamma_k, \delta_k\}$:

$$\begin{aligned}\hat{\mathbf{x}}_k &= \mathbb{E}[\mathbf{x}_k \mid Z_k] \\ &= \sum_{\gamma,\delta} \mathbb{E}[\mathbf{x}_k \mid Z_k, \gamma_k = \gamma, \delta_k = \delta] \Pr\{\gamma_k = \gamma, \delta_k = \delta \mid Z_k\}.\end{aligned} \tag{3}$$

Some of the conditional expectations in (3) may be computed precisely and some may be reasonably approximated. For $\gamma_k = 0$ the current measurement does not carry valuable information about the state. It thus follows that, for any $\delta \in \{1, C\}$, $\mathbb{E}[\mathbf{x}_k \mid Z_k, \gamma_k = 0, \delta_k = \delta] = A\hat{\mathbf{x}}_{k-1}$ where we have utilized the fact that $\{\mathbf{z}_k, \gamma_k = 0, \delta_k = \delta\}$ is independent of $\{\mathbf{x}_{k-1}, Z_{k-1}\}$. In other words, given that the last measurement is faulty, it does not carry useful information about the state at the previous time. Thus, the optimal estimate of $\mathbf{x}_k$ given that the last measurement is faulty is just a linear extrapolation of the previously obtained estimate.

The remaining terms are $\mathbb{E}[\mathbf{x}_k \mid Z_k, \gamma_k = 1, \delta_k = 1]$ and $\mathbb{E}[\mathbf{x}_k \mid Z_k, \gamma_k = 1, \delta_k = C]$. We interpret these as optimal estimates of $\mathbf{x}_k$ given that the latest observation is informative, that is multiplicative fault-free (albeit noisier than a nominal one). We approximate them by considering the following suboptimal reasoning. Instead of looking for overall optimality, we shall approximate the term $\mathbb{E}[\mathbf{x}_k|Z_k, \gamma_k = 1, \delta_k = \delta]$ by a step-wise near-optimal expression. Namely, we shall compute the estimate $\mathbf{x}_k$ that is near-optimal with respect to the previously obtained $\hat{\mathbf{x}}_{k-1}$ and the current, fault-free measurement $\mathbf{z}_k$. Mathematically, this approximation reads,

$$\mathbb{E}[\mathbf{x}_k|Z_k, \gamma_k = 1, \delta_k = \delta] \approx \mathbb{E}[\mathbf{x}_k|\hat{\mathbf{x}}_{k-1}, \mathbf{z}_k, \gamma_k = 1, \delta_k = \delta]. \tag{4}$$

Note, that in the standard (fault-free) Kalman filter (KF) setting (4) holds with equality, since the information carried by the entire measurement sequence is equivalent to that carried by the latest estimate $\hat{\mathbf{x}}_{k-1}$ and the the new measurement $\mathbf{z}_k$. A similar approximation is used in the GPB filters, where the whole history is merged into a single estimate from the previous step.

Using (4), an estimate may be obtained via a usual KF update step that uses the previously obtained estimate together with the latest measurement to yield the next estimate. Formally,

$$\mathbb{E}[\mathbf{x}_k \mid \hat{\mathbf{x}}_{k-1}, \mathbf{z}_k, \gamma_k = 1, \delta_k = \delta] \approx F_1\hat{\mathbf{x}}_{k-1} + F_2\mathbf{z}_k, \tag{5}$$

where $F_1 = (I - F_2H)A$, and

$$F_2 = (AP_{k-1}A^T + Q)H^T(H(AP_{k-1}A^T + Q)H^T + R\delta)^{-1} \tag{6}$$

where P_k is the estimation error covariance, given by the recursion

$$P_k = (I - F_2(k-1)H)(AP_{k-1}A^T + Q). \tag{7}$$

Consider next the weighting factors in (3)

$$\begin{aligned}\mu_{\gamma,\delta}(k) &\triangleq \Pr\{\gamma_k = \gamma, \delta_k = \delta \mid Z_k\} \\ &= \frac{\mathrm{p}(\mathbf{z}_k \mid Z_{k-1}, \gamma_k = \gamma, \delta_k = \delta)}{\mathrm{p}(\mathbf{z}_k \mid Z_{k-1})} \Pr\{\gamma_k = \gamma, \delta_k = \delta \mid Z_{k-1}\}\end{aligned} \tag{8}$$

For $\gamma = 1$, the term in the numerator is the likelihood of the present measurement conditioned on the entire history as well as on the event that it is not a faulty measurement. If all past measurements were fault-free, the above density would be Gaussian with mean $\hat{\mathbf{z}}_k = H\hat{\mathbf{x}}_k$ and (innovation) covariance $S_k = HP_kH^T + R$, that may be easily obtained from the standard KF procedure. Since this is not the case, the following is an approximation:

$$\mathrm{p}(\mathbf{z}_k \mid Z_{k-1}, \gamma_k = 1, \delta_k = \delta) \approx \mathcal{N}(\mathbf{z}_k; \hat{\mathbf{z}}_k, HP_kH^T + R\delta). \tag{9}$$

The likelihood of $\mathbf{z}_k$ given that it is a faulty measurement, i.e. $\gamma = 0$, may be computed precisely, as in this case $\mathbf{z}_k$ is independent of the history $\mathrm{p}(\mathbf{z}_k \mid Z_{k-1}, \gamma_k = 0, \delta_k = \delta) = \mathcal{N}(\mathbf{z}_k; 0, R\delta)$. The last term on the RHS of (8) may be expanded using the total probability law as

$$\begin{aligned}&\Pr\{\gamma_k = \gamma, \delta_k = \delta \mid Z_{k-1}\} \\ &\quad = \sum_{\gamma_1,\delta_1} \Pr\{\gamma_k = \gamma, \delta_k = \delta \mid Z_{k-1}, \gamma_{k-1} = \gamma_1, \delta_{k-1} = \delta_1\} \\ &\qquad\qquad \times \mu_{\gamma_1,\delta_1}(k-1).\end{aligned} \tag{10}$$

Unlike previously considered terms, the remaining conditional probability is affected by the differences between various problems posed at the beginning of this section. We treat each case independently.

Problem 1. In this case the vector valued sequence $\{\gamma_k, \delta_k\}$ is i.i.d. Hence,

$$\Pr\{\gamma_k = \gamma, \delta_k = \delta \mid Z_{k-1}\} = \Pr\{\gamma_k = \gamma\}\Pr\{\delta_k = \delta\}. \quad (11)$$

Problem 2. Similarly to the case of Problem 1 we have

$$\Pr\{\gamma_k = \gamma, \delta_k = \delta \mid Z_{k-1}\} = \Pr\{\gamma_k = \gamma, \delta_k = \delta\}. \quad (12)$$

Problem 3. The derivation of the filter for this problem rests on the following lemma.

Lemma 1. *Let $\{a_k\}$ and $\{b_k\}$ be independent Markov sequences. Then, the vector-valued sequence $\{a_k, b_k\}$ is Markov.*

Equipped with Lemma 1, the derivation of the filter for Problem 3 boils down to the following relation:

$$\begin{aligned}&\Pr\{\gamma_k = \gamma, \delta_k = \delta \mid Z_{k-1}, \gamma_{k-1} = \gamma_1, \delta_{k-1} = \delta_1\} \\ &\quad = \Pr\{\gamma_k = \gamma \mid \gamma_{k-1} = \gamma_1\}\Pr\{\delta_k = \delta \mid \delta_{k-1} = \delta_1\}. \quad (13)\end{aligned}$$

Problem 4. The solution to Problem 4 is obtained in a similar manner to Problem 3. The only difference is that the joint conditional probability in the last equation cannot be decomposed into a product of marginal conditional probabilities.

$$\begin{aligned}&\Pr\{\gamma_k = \gamma, \delta_k = \delta \mid Z_{k-1}, \gamma_{k-1} = \gamma_1, \delta_{k-1} = \delta_1\} \\ &\quad = \Pr\{\gamma_k = \gamma, \delta_k = \delta \mid \gamma_{k-1} = \gamma_1, \delta_{k-1} = \delta_1\}. \quad (14)\end{aligned}$$

We have, thus, obtained efficient estimation schemes for Problems 1–4, that are based on a step-wise approximation of the conditional expectation. In all cases, three primitive KFs, each matched to a different joint fault indicator event, run in parallel and produce local, conditional estimates. These estimates are then fused into a single output (to be used as the filter's input in the next iteration) using weighting probabilities that are computed using the outputs of the KFs. The differences between different problems are in the computation of these weighting probabilities and they stem from the different mechanisms controlling the transitions between faults at different times.

In the next section we adopt an alternative, suboptimal approach for estimating the state for the posed problems. This approach is based on the well known IMM method, and uses the same number of primitive filters with different fusion mechanism.

IV. IMM-Based Filters

To invoke the IMM algorithm we consider the posterior density of the state $\mathbf{x}_k$ given all the measurements $Z_k = \{\mathbf{z}_0, ..., \mathbf{z}_k\}$.

$$\begin{aligned}&\mathrm{p}(\mathbf{x}_k \mid Z_k) \\ &= \sum_{\gamma,\delta} \mathrm{p}(\mathbf{x}_k \mid Z_k, \gamma_k = \gamma, \delta_k = \delta)\Pr\{\gamma_k = \gamma, \delta_k = \delta \mid Z_k\}.\end{aligned} \quad (15)$$

The weighting probabilities may be computed in the same manner as in Section III. The difference is introduced in the approximate calculation of the conditional densities carrying the information of the (conditional) estimates $\mathbb{E}[\mathbf{x}_k \mid Z_k, \gamma_k = \gamma, \delta_k = \delta]$. We outline this calculation next.

Consider the conditional density in (15). We may rewrite it using Bayes' rule as follows

$$\begin{aligned}&\mathrm{p}(\mathbf{x}_k \mid Z_k, \gamma_k = \gamma, \delta_k = \delta) \\ &= \frac{\mathrm{p}(\mathbf{z}_k \mid \gamma_k = \gamma, \delta_k = \delta, \mathbf{x}_k)}{\mathrm{p}(\mathbf{z}_k \mid \gamma_k = \gamma, \delta_k = \delta, Z_{k-1})}\mathrm{p}(\mathbf{x}_k \mid Z_{k-1}, \gamma_k = \gamma, \delta_k = \delta).\end{aligned} \quad (16)$$

This equation represents the measurement update step of a standard KF matched to the event $\{\gamma_k = \gamma, \delta_k = \delta\}$ which represents the active model at time k.

The IMM technique is utilized in the approximate computation of the corresponding time update step, represented by the conditional prior density $\mathrm{p}(\mathbf{x}_k \mid Z_{k-1}, \gamma_k = \gamma, \delta_k = \delta)$. To this end, we define three possible modes, as follows: $m_1 \triangleq \{\gamma_k = 1, \delta_k = 1\}$, $m_2 \triangleq \{\gamma_k = 1, \delta_k = C\}$, $m_3 \triangleq \{\gamma_k = 0\}$. Consequently, three primitive KFs, each matched to a different mode, are run in parallel. At each estimation cycle, all three filters are fed by the current measurement as well as by new initial conditions. The initial conditions are obtained from an interaction of all three filters using their outputs as inputs to the interaction block. Thus, poorly performing filters are "punished", whereas well-performing ones are "rewarded" in terms of their initial conditions. The overall output is obtained, at each step, as a weighted combination of the individual outputs of the mode-matched filters.

The steps of the IMM algorithm are summarized below:

1) Mixing probabilities computation:
$$\begin{aligned}\mu_{i|j}(k-1 \mid k-1) &= \Pr\{\mathcal{M}_{k-1} = m_i \mid \mathcal{M}_k = m_j, Z_{k-1}\} \\ &= \frac{1}{c} p_{ij}\mu_i(k-1),\ i, j = 1, 2, 3, \end{aligned} \quad (17)$$
where $p_{ij} \triangleq \Pr\{\mathcal{M}_k = m_j \mid \mathcal{M}_{k-1} = m_i\}$ and c is some normalization constant that may depend on j.
2) Mixing: starting with the estimate matched to model m_i from time $k-1$, $\hat{\mathbf{x}}^i_{k-1}$, compute the mixed initial condition for the filter matched to m_j:
$$\hat{\mathbf{x}}^{0j}(k-1) = \sum_{i=1}^{3} \hat{\mathbf{x}}^i_{k-1}\mu_{i|j}(k-1 \mid k-1),\ j = 1, 2, 3 \quad (18)$$
and the associated covariance P^{0j}_{k-1}.
3) Mode-matched filtering: using the initial conditions of step 2 as inputs to the filter matched to m_j, obtain the mode-matched estimate $\hat{\mathbf{x}}^j_k$ and P^j_k, as well as the associated likelihood function $\Lambda_j(k) \triangleq \mathrm{p}(\mathbf{z}_k \mid \mathcal{M}_k = m_j, Z_{k-1})$ for $j = 1, 2, 3$.
4) Mode probability update: the probabilities $\mu_j(k),\ j = 1, 2, 3$ are calculated similarly to $\mu_{\gamma,\delta}(k)$ in (8).
5) Estimate combination: $\hat{\mathbf{x}}(k) = \sum_{j=1}^{3} \hat{\mathbf{x}}^j_k\mu_j(k)$, with the corresponding covariance.

Note that the filter matched to m_3 assumes that no measurement carries useful information. Therefore, it simply propagates its initial conditions over time.

V. Linear Optimal Filters

Next we derive the linear optimal filters for the different cases of fault sequences posed in Section II. We consider the following recursive form for $\hat{\mathbf{x}}_k \triangleq \hat{\mathbf{x}}_{k|k}$:

$$\hat{\mathbf{x}}_{k+1} = F_1(k)\hat{\mathbf{x}}_k + F_2(k)\mathbf{z}_{k+1}. \quad (19)$$

For $\hat{\mathbf{x}}_k$ to be MMSE-optimal, it should satisfy the following necessary and sufficient orthogonality conditions [9]:

$$\mathbb{E}\left[(\hat{\mathbf{x}}_{k+1} - \mathbf{x}_{k+1})\mathbf{z}_i^T\right] = 0, \quad i = 0, ..., k+1. \quad (20)$$

We utilize these conditions for computing $F_1(k)$ and $F_2(k)$. We proceed similarly to [1], [3], omitting the dependence of F_1 and F_2 on k, as well as the dependence of $\mathbf{v}_{k+1}$ on δ_{k+1}, for brevity.

A. Filter Derivation

Consider the orthogonality conditions (20) for $i = 0, ..., k$. Substituting (1) and (19) and utilizing the independence of $\mathbf{w}_k$ and $\{\gamma_{k+1}, \mathbf{z}_i\}$, it can be shown that the following relations guarantee satisfaction of the orthogonality conditions for $i = 0, ..., k$

$$F_1 = A - F_2 H A \Pr\{\gamma_{k+1} = 1 \mid \gamma_i = 1\}. \tag{21}$$

Clearly, (21) holds iff its RHS does not depend on i. Thus

Lemma 2. *Sufficient conditions for the first $k+1$ orthogonality conditions of* (20) *to hold are that* $\Pr\{\gamma_{k+1} = 1 \mid \gamma_i = 1\}$ *be independent of i for $i = 0, \ldots, k$.*

A similar lemma was originally presented in [3] for the case of multiplicative faults, where the above conditions were shown to be necessary and sufficient. However, for the necessity, H was required to be of rank n (the dimension of $\mathbf{x}_k$) which is a very strong requirement (since it means that all components of the state are observed through the measurement equation).

Note that for the fault-free case, where $\Pr\{\gamma_i = 1\} = 1$, we obtain the standard KF equation as expected.

Corollary 1. *Let $\{\gamma_k\}$ be a homogeneous Bernoulli Markov chain satisfying the sufficient conditions of Lemma 2, and suppose that not all elements of the transition matrix of $\{\gamma_k\}$ are in $\{0, 1\}$. Then one of the following conditions holds*

1) $\Pr\{\gamma_{k+1} = 1 \mid \gamma_k = 1\} = 1.$
2) *The transition probability matrix of $\{\gamma_k\}$ has identical rows.*

When either of the conditions of Corollary 1 holds, we shall denote the transition probability $\Pr\{\gamma_{k+1} = 1 \mid \gamma_i = 1\} = \Pr\{\gamma_{k+1} = 1 \mid \gamma_k = 1\}$ as p_{11}.

To obtain the second relation between F_1 and F_2 we consider (20) for $i = k+1$. Taking the expectation we have:

$$\mathbb{E}[\gamma_{k+1}](I - F_2 H) C_{k+1} H^T - F_1 p_{11} \mathbb{E}\left[\hat{\mathbf{x}}_k \mathbf{x}_{k+1}^T\right] H^T - F_2 R \mathbb{E}[\delta_{k+1}] = 0, \tag{22}$$

where $C_k = \mathbb{E}\left[\mathbf{x}_k \mathbf{x}_k^T\right]$.

For convenience we summarize the filter equations below.

$$\hat{\mathbf{x}}_{k+1} = F_1 \hat{\mathbf{x}}_k + F_2 \mathbf{z}_{k+1} \tag{23a}$$

$$F_1 = A - p_{11} F_2 H A \tag{23b}$$

$$F_2(k) = \left(\mathbb{E}[\gamma_{k+1}] C_{k+1} - p_{11} A V_k A^T\right) H^T \times \left[\mathbb{E}[\gamma_{k+1}] H C_{k+1} H^T - p_{11}^2 H A V_k A^T H^T + R \mathbb{E}[\delta_{k+1}]\right]^{-1} \tag{23c}$$

$$C_{k+1} = A C_k A^T + Q \tag{23d}$$

$$V_{k+1} = F_1 V_k F_1^T + F_2(\mathbb{E}[\gamma_{k+1}] H C_{k+1} H^T + R \mathbb{E}[\delta_{k+1}]) F_2^T + p_{11} F_2 H A V_k F_1^T + p_{11} F_1 V_k A^T H^T F_2^T \tag{23e}$$

B. Special Cases

Equations (23) are general. The terms affected by the specific problem under consideration are p_{11} and $\mathbb{E}[\gamma_{k+1}] = \Pr\{\gamma_{k+1} = 1\}$. Their computation is discussed next.

Problem 1. In this case $p_{11} = \Pr\{\gamma_{k+1} = 1\} = \mathbb{E}[\gamma_{k+1}]$. In the additive fault-free case, i.e., $\Pr\{\delta_k = 1\} = 1$, this reduces to the classical result of Nahi [1].

Problem 2. It follows that $\{\gamma_k\}$ is an i.i.d. sequence as well. Thus,

$$\Pr\{\gamma_{k+1} = 1 \mid \gamma_i = 1\} = \Pr\{\gamma_{k+1} = 1, \delta_{k+1} = 1\} + \Pr\{\gamma_{k+1} = 1, \delta_{k+1} = 0\}. \tag{24}$$

Problem 3. In this case, following Corollary 1, p_{11} is assumed known. In the additive fault-free case, this filter reduces to that of [3].

Problem 4. Similarly to the case of Problem 3, the desired quantities may be obtained from the joint law of $\{\gamma_k\}$ and $\{\delta_k\}$.

VI. Simulation Study

In this section we compare the performance of the optimal linear algorithm with that of the one-step near-optimal nonlinear filter in a simple scalar example.

A. General Simulation Setup

Consider the following system, adapted from the example in [3].

$$\mathbf{x}_{k+1} = -0.8 \mathbf{x}_k + \mathbf{w}_k, \tag{25}$$

where $\mathbf{x}_k$ is a scalar state variable, $\mathbf{w}_k$ is a zero-mean, unit variance, white Gaussian noise and $\mathbf{x}_0 = 0$. The system is observed via one of two channels. The first channel introduces only additive faults:

$$\mathbf{z}_k = \mathbf{x}_k + \mathbf{v}_k(\delta_k). \tag{26}$$

The second channel is subject to both kinds of faults:

$$\mathbf{z}_k = \beta_k \mathbf{x}_k + \mathbf{v}_k(\delta_k). \tag{27}$$

The sequence $\{\mathbf{v}_k\}$ is a zero mean white Gaussian process independent of $\{\mathbf{w}_k\}$, and $\{\beta_k\}$ is a Bernoulli Markov fault sequence of the second channel with known transition probability matrix P.

The first channel is the actual measurement channel with probability $1-q$, and the second one is chosen with probability q. The choice between the two channels is done independently of $\{\mathbf{w}_k\}$ and $\{\mathbf{v}_k\}$. The sequence of the additive fault variables $\{\delta_k\}$ is chosen to be i.i.d. such that each δ_k assumes the values 1 and C with probabilities r and $1-r$, respectively.

Both measurement channels constitute a unique observation equation of the following form

$$\mathbf{z}_k = \gamma_k \mathbf{x}_k + \mathbf{v}_k(\delta_k), \tag{28}$$

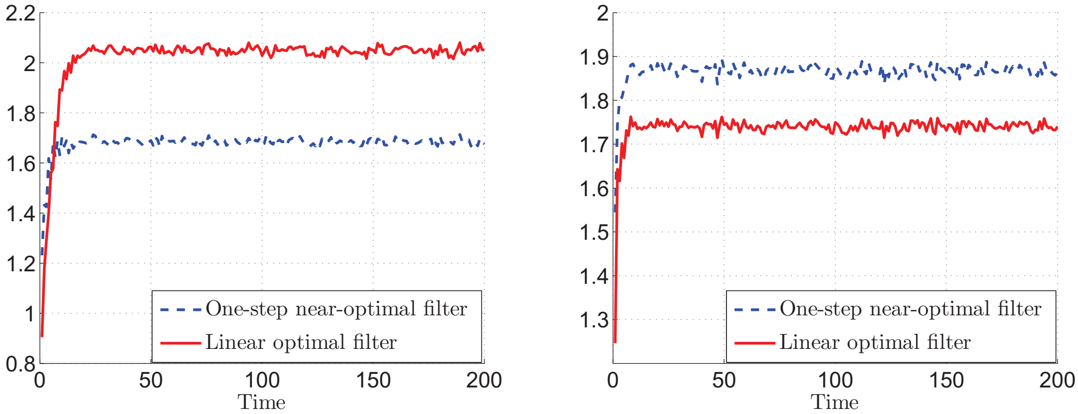
where $\gamma_k = 1 - \alpha + \alpha \beta_k$, and $\Pr\{\alpha = 1\} = 1 - \Pr\{\alpha = 0\} = q$. The formulation (28) induces various transition structures of $\{\gamma_k\}$, as discussed in the next subsection.

In all examples in the sequel the following common parameters are used $C = 10$, $R = 1$, $r = 0.8$, $p = 0.9$.

B. Example 1: Markov Transitions

In this experiment we set $q = 1$ and $P = \begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix}$. Consequently, $\gamma_k = \beta_k$, meaning that the sequence of multiplicative faults indicators is a Markov process. Since the sequence $\{\delta_k\}$ is i.i.d. and, in particular, a Markov process, the conditions of Problem 3 hold. Note that in this case the necessary conditions of Corollary 1 are not satisfied for $p \notin \{0, \frac{1}{2}\}$. Hence, the sufficient conditions of Lemma 2 do not hold, and no claim can be made regarding the optimality of the linear estimator.

We compare the mean squared error of this linear filter to that of the one step near-optimal filter devised in Section III. As an overall reference we also generate the optimal estimate using a "genie-based" KF that knows the exact values of all fault indicators at every time. Knowing the faults at each time degenerates the problem to the standard setting such that a standard KF yields optimal results.



(a) Example 1 **(b) Example 2**

Figure 1: Normalized estimation errors of the linear optimal filter (dashed), and the one-step optimal filter (solid)

Both error sequences, normalized by the errors of the genie-based filter and averaged over 10000 Monte Carlo runs are presented, versus time, in Fig. 1(a). Note that in this case, whereas the linear filter acts under conditions it is not designed for, the nonlinear estimator operates under Markovian dynamics of the modes it was designed for. It is, therefore, not surprising that the errors produced by the linear filter are higher than those of the nonlinear one.

C. Example 2: Non-Markov Transitions

In this example we deviate slightly from the scope of the problems defined in Section II by considering the system (25)–(28) such that the 2nd condition of Corollary 1 is satisfied, and the linear filter devised earlier is the linear, MMSE-optimal estimator for the problem at hand. Since a Markov process satisfying the above conditions has a somewhat degenerate structure, we consider a more complicated case under which the Markov property is not valid. To this end we set $q = 0.7$ and $P = \left(\begin{smallmatrix} 1-p & p \\ 1-p & p \end{smallmatrix}\right)$, such that $\gamma_k \neq \beta_k$. It can be shown [3] that the transition probability $\Pr\{\gamma_k = 1 \mid \gamma_j = i\}$ has the following form:

$$\Pr\{\gamma_k = 1 \mid \gamma_j = i\} = \begin{cases} p, & i = 0 \\ \frac{1-q(1-p^2)}{1-q(1-p)}, & i = 1 \end{cases}. \qquad (29)$$

Since $\Pr\{\gamma_k = 1 \mid \gamma_j = 0\} \neq \Pr\{\gamma_k = 1 \mid \gamma_j = 1\}$, and $p \neq 1$ and $q \neq 1$, neither the rows of the corresponding transition matrix are identical, nor $\Pr\{\gamma_k = 1 \mid \gamma_{k-1}\} = 1$ and, according to Corollary 1, the process is not Markov.

We compare the errors (normalized by those generated by the genie-based filter) of the linear and the nonlinear filters. The results are presented in Fig. 1(b).

In this case, the linear filter is optimal within the family of all linear filters, but the one-step optimal filter operates under conditions deviating significantly from those it is designed for since the mode transitions are no longer Markov and the transition kernel of Eq. (29) is only an approximation of a TPM that is required in the one-step optimal filter. Not surprisingly, the performance of the optimal linear filter is superior to that of the nonlinear one.

VII. Concluding Remarks

We have presented a unified formulation of a general fault-prone dynamical system where faults affect the measurement equation. Unlike previously reported research, the proposed formulation admits simultaneous presence of both multiplicative faults representing interruptions indicators, and additive faults representing measurement biases. Both types of faults may be decoupled, or they can be strongly correlated, such that classical problems are obtained as special cases when the proposed formulation is appropriately degenerated. Several special problems resulting from the general formulation have been listed, and three families of suboptimal state estimators have been derived for these problems. These algorithm families include 1) one-step near-optimal filters, 2) IMM-based filters, and 3) recursive linear optimal filters. The linear optimal and one-step near-optimal filters have been tested in simulation, demonstrating their relative performance and providing rules of thumb on the conditions when either of the approaches should be preferred.

References

[1] N. Nahi, "Optimal recursive estimation with uncertain observation," *IEEE Trans. on Information Theory*, vol. IT-15, no. 4, pp. 457–462, July 1969.
[2] Y. Sawaragi, T. Katayama, and S. Fujishige, "Adaptive estimation for a linear system with interrupted observation," *IEEE Trans. on Automatic Control*, vol. 18, no. 2, pp. 152–154, 1973.
[3] M. Hadidi and S. Schwartz, "Linear recursive state estimators under uncertain observations," *IEEE Trans. on Automatic Control*, vol. AC-24, no. 6, pp. 944–948, December 1979.
[4] O. Costa, "Linear minimum mean square error estimation for discrete-time Markovian jump linear systems," *IEEE Trans. on Automatic Control*, vol. 39, no. 8, pp. 1685–1689, 1994.
[5] I. Rapoport and Y. Oshman, "Efficient fault tolerant estimation using the IMM methodology," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 43, no. 2, pp. 492–508, 2007.
[6] ——, "A new estimation error lower bound for interruption indicators in systems with uncertain measurements," *IEEE Trans. on Information Theory*, vol. 50, no. 12, pp. 3375–3384, 2004.
[7] G. Ackerson and K. Fu, "On state estimation in switching environments," *IEEE Trans. on Automatic Control*, vol. 15, no. 1, pp. 10–17, 1970.
[8] H. Blom and Y. Bar-Shalom, "The interacting multiple model algorithm for systems with Markovian switching coefficients," *IEEE Trans. on Automatic Control*, vol. 33, no. 8, pp. 780–783, 1988.
[9] A. Jazwinski, *Stochastic Processes and Filtering Theory*. Academic Press, 1970.