Analysis of LTV Fault Detection Schemes with Additive Faults

Timothy J. Wheeler, Peter Seiler, Andrew K. Packard, and Gary J. Balas

Abstract—This paper considers the problem of certifying the performance of a class of model-based fault detection schemes. The underlying plant is assumed to be a linear time-varying (LTV) system subject to a Markov-switching fault input. The fault detection scheme consists of two parts: an LTV component that produces a scalar residual and a static nonlinear function that infers the presence of a fault based on this residual. Probabilistic performance metrics are presented and the complexity of computing these metrics is analyzed. It is shown that under a set of realistic assumptions, this complexity is reduced to polynomial time. An aerospace example, involving a pitot-static probe subject to random bias faults, is used to demonstrate the usefulness of this analysis.

I. INTRODUCTION

In safety critical applications, a system must not only be highly reliable, but that reliability must be certiﬁable. This is particularly true in civil aviation, where the FAA requires fly-by-wire control systems to have fewer than $10^{-9}$ catastrophic failures per flight-hour [1]. Such system-wide failures can occur if the system is rendered inoperable by a critical component failure or if the system performs poorly because of an undetected component failure. One approach, commonly found in avionics systems, is to use parallel redundant components, which ensures the availability of the system, even in the presence of component failures [1], [2], [3]. A failed component is detected by directly comparing the behavior of each redundant component. Hence, these schemes tend to detect faults accurately, and their performance is simple to certify using fault trees [4].

However, in some applications, such as Unmanned Aerial Vehicles (UAVs), the designer cannot afford the extra weight, size, and power needed to support identical redundant components. To prevent system-wide failures due to an undetected component failure, the analytical redundancies between components can be exploited to detect faults. For example, if three measurements $m_1$, $m_2$, and $m_3$ are available, and these quantities are known to satisfy the analytical relations $m_1 = f_1(m_2, m_3)$, $m_2 = f_2(m_1, m_3)$, and $m_3 = f_3(m_1, m_2)$, then the residuals $r_1 = m_1 - f_1(m_2, m_3)$, $r_2 = m_2 - f_2(m_1, m_3)$, etc. can be used to detect failures in the components that produce these measurements. This approach certainly reduces the number of individual components needed; however, there are two main drawbacks to consider. First, merely identifying a fault cannot prevent system-wide failure if the failed component is indispensable (i.e. no other components can perform the same critical function). Second, the performance of fault detection schemes based on analytical redundancy can be difficult to certify if the analytical relationships are dynamic or nonlinear. While the first difficulty is unavoidable, this paper addresses the second difficulty.

Although there is a vast body of literature on model-based fault detection and identification (FDI) (e.g., [5], [6], [7]), little attention is given to the rigorous evaluation of the performance and reliability metrics required to certify safety-critical aerospace systems. Monte Carlo methods [8] provide a statistically rigorous approach to performance analysis, but it can difficult to quantify the error present in the results. The goal of this paper is to define a set of performance metrics that can be computed analytically and to provide a class of systems for which these metrics can be computed efficiently.

The problem of fault detection is modeled by the interconnection of systems shown in Fig. 1. Section II provides a detailed description of the plant $P$, the fault detection scheme $(F, \delta)$, and the associated input and output signals. The framework presented here is a generalization of the fault detection problem considered in [9]. To quantify the performance of a fault detection scheme that falls within this framework, a set of performance metrics is defined in Section II-C. The analysis in Section III demonstrates that, in general, computing these metrics is an intractable problem, but for a specific class of fault detection problems these metrics can be computed in polynomial time. Section III-C presents conditions under which this complexity can be further reduced to quadratic or linear time. In Section IV, this analysis is applied to a fault detection problem involving a pitot-static probe subject to randomly occurring bias faults. Finally, Section V discusses the conclusions of this work, as well as avenues of future research.

II. PROBLEM STATEMENT

First, we establish some notation. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the underlying probability space, and let $\mathcal{K} := \{k \in \mathbb{Z} | k \geq 0\}$ be the discrete time index set. For any stochastic process $v : \Omega \times \mathcal{K} \to \mathbb{R}^n$, the notation $\{v_k\}_{k \in \mathcal{K}}$ or $\{v_k\}$ represents the entire process. For each $k \in \mathcal{K}$, the notation $v_k$ represents the random variable $v_k : \omega \to v_k(\omega)$. For $i, j \in \mathcal{K}$ with $i \leq j$, define the partial sequence $\{v_{i:\j} := \{v_i, v_{i+1}, \ldots, v_j\}\}$.

A. Plant Model

Assume that the plant, labeled $P$ in Fig. 1, is of the form

\begin{equation}
\begin{align*}
x_{k+1} &= \hat{A}_k x_k + \hat{B}_{u,k} u_k + \hat{B}_{v,k} v_k + f_k(\theta_{u,k}) \\
y_k &= \hat{C}_k x_k + \hat{D}_{u,k} u_k + \hat{D}_{v,k} v_k + f_k(\theta_{v,k}),
\end{align*}
\end{equation}

T. J. Wheeler and A. K. Packard are with the Dept. of Mechanical Engineering, University of California, Berkeley, CA 94708.
P. Seiler and G. J. Balas are with the Dept. of Aerospace Engineering & Mechanics, University of Minnesota, Minneapolis, MN 55455.

Corresponding author: T. J. Wheeler (t wheeler@berkeley.edu).
The role of the system $F$ is in the fault detection scheme $(F, \delta)$, and the output $d$ is the decision or inference made by $(F, \delta)$.

where $\{u_k\}$ is a known deterministic input and $\{v_k\}$ is an i.i.d., Gaussian stochastic process with $v_k \sim \mathcal{N}(0, I)$, for all $k$. The random occurrence of faults is modeled by a time-homogeneous Markov chain $\{\theta_k\}$ with a finite state space $\mathcal{M} := \{0, 1, \ldots, m\}$, transition probability matrix

$$
\Pi_{ij} := \mathbb{P}(\theta_{k+1} = j \mid \theta_k = i), \quad i, j \in \mathcal{M},
$$

and initial distribution $\pi_0$ [10]. The random variable $\theta_k$ is called the mode of the system at time $k$. The signal $\{f_k\}$ is a sequence of deterministic functions that map the mode sequence $\theta_{0:k}$ to an additive fault input $f_k(\theta_{0:k})$, such that for all $k$, $\theta_k = 0$ implies $f_k(\theta_{0:k}) = 0$. That is, the event $\{\theta_k = 0\}$ occurs when the system is in the nominal mode (i.e., no faults) at time $k$. Conditional on the event $\{\theta_{0:k} = \theta_{0:k}\}$, the system (1) is a linear time-varying (LTV) Gaussian system driven by the known deterministic inputs $\{u_k\}$ and $\{f_k(\theta_{0:k})\}$ and the random input $\{v_k\}$. Hence, the system (1) belongs to the class of conditionally-linear-Gaussian systems [11].

### B. Fault Detection Scheme

Since the random occurrence of faults is modeled by the Markov chain $\{\theta_k\}$, the general purpose of fault detection and identification is to infer something about the current value of $\{\theta_k\}$. More precisely, if for some $s$ the set $\mathcal{M}$ is partitioned into $s$ disjoint subsets $\mathcal{M}_0, \mathcal{M}_1, \ldots, \mathcal{M}_{s-1}$ and we define $\mathcal{D} = \{0, 1, \ldots, s-1\}$, then the goal is to determine for which $d \in \mathcal{D}$ is the event $\{\theta_k \in \mathcal{M}_d\}$ most likely at time $k$. In this paper, we consider the simplest case, called fault detection, where $\mathcal{M}_0 := \{0\}$, $\mathcal{M}_1 := \{1, 2, \ldots, m\}$, and $\mathcal{D} := \{0, 1\}$. In other words, the objective is to determine if the system is in the nominal mode or in some fault mode.

Assume that fault detection scheme, labeled $F$ and $\delta$ in Fig. 1, is modeled as a deterministic LTV system $F$ given by

$$
\begin{align*}
    x_{k+1} &= A\hat{x}_k + B_{u,k}u_k + B_{v,k}v_k, \\
    y_k &= C\hat{x}_k + D_{u,k}u_k + D_{v,k}v_k,
\end{align*}
$$

and a sequence of static, memoryless, deterministic functions $\{\delta_k\}$, where each $\delta_k : \mathbb{R} \rightarrow \mathcal{D}$ is called a decision function. The output $\{d_k\}$, given by $d_k = \delta_k(r_k) \in \mathcal{D}$, indicates the decision or inference made by the fault detection scheme. The role of the system $F$ is to generate an output $\{r_k\}$, known as the residual, that has small mean and variance when $\{\theta_k\}$ is in $\mathcal{M}_0$ and has large mean and large variance otherwise. The complementary role of each decision function $\delta_k$ is to determine when the residual $r_k$ is large enough to indicate that the event $\{\theta_k \in \mathcal{M}_1\}$ is likely at time $k$. A commonly-used decision function is

$$
\delta(r) = \mathbb{I}(|r| > \varepsilon),
$$

where $\mathbb{I}$ is the indicator function and $\varepsilon > 0$ is the threshold.

### C. Probabilistic Analysis

For each $k \in \mathcal{K}$, define the events

$$
\begin{align*}
    H_{0,k} &:= \{\theta_k \in \mathcal{M}_0\}, \quad R_{0,k} := \{d_k = 0\}, \\
    H_{1,k} &:= \{\theta_k \in \mathcal{M}_1\}, \quad R_{1,k} := \{d_k = 1\}.
\end{align*}
$$

The reliability of the plant is given by $H_{0,k}$ and $H_{1,k}$, while the behavior of the fault detection scheme is given by $R_{0,k}$ and $R_{1,k}$ [12]. Note that $\{H_{0,k}, H_{1,k}\}$ and $\{R_{0,k}, R_{1,k}\}$ both form partitions of the sample space $\Omega$. For each $k \in \mathcal{K}$, the performance of the fault detection scheme, with respect to the plant, is given by the following four events: a true negative, $R_{0,k} \cap H_{0,k}$; a false positive, $R_{1,k} \cap H_{0,k}$; a false negative, $R_{0,k} \cap H_{1,k}$; and a true positive, $R_{1,k} \cap H_{1,k}$ [12]. These events also form a partition of the sample space $\Omega$, and their corresponding probabilities are denoted

$$
\begin{align*}
    P_k^{TN} &= \mathbb{P}(R_{0,k} \cap H_{0,k}), \\
    P_k^{FP} &= \mathbb{P}(R_{1,k} \cap H_{0,k}), \\
    P_k^{FN} &= \mathbb{P}(R_{0,k} \cap H_{1,k}), \\
    P_k^{TP} &= \mathbb{P}(R_{1,k} \cap H_{1,k}).
\end{align*}
$$

These probabilities provide all the necessary information, because their values can be used to compute the probability of any event in the $\sigma$-algebra generated by the collection $\{R_{0,k}, R_{1,k}, H_{0,k}, H_{1,k}\}$. Since the values of (3)–(6) sum to one, only three of the four quantities must be computed.

Although the probabilities (3)–(6) provide all the necessary information, their numerical values can be difficult to interpret. For example, suppose that $\mathbb{P}(H_{1,k}) \approx 0$ for $k = 0, 1, \ldots, T$. This implies that

$$
\mathbb{P}(H_{1,k}) = P_k^{FN} + P_k^{TP} \approx 0.
$$

Since both $P_k^{FN}$ and $P_k^{TP}$ are small, it is difficult to get a sense of how well the fault detection scheme will perform in the presence of a fault at times $k \in \{0, 1, \ldots, T\}$. In this case, it is beneficial to consider the relative magnitudes of $P_k^{FN}$ and $P_k^{TP}$. This approach gives rise to two conditional probabilities: the probability of detection

$$
P_k^D := \mathbb{P}(R_{1,k} \mid H_{1,k}) = \frac{P_k^{TP}}{P_k^{TN} + P_k^{TP}},
$$

and the probability of a false alarm

$$
P_k^F := \mathbb{P}(R_{1,k} \mid H_{0,k}) = \frac{P_k^{TP}}{P_k^{TN} + P_k^{TP}}.
$$

If the probability $\mathbb{P}(H_{1,k})$ is known, equations (7) and (8) can be rearranged to compute the probabilities (3)–(6) from $P_k^D$, $P_k^F$, and $\mathbb{P}(H_{1,k})$.

### III. COMPUTATIONAL ISSUES

For each $k \in \mathcal{K}$, the joint density of the residual $r_k$ and the mode $\theta_k$ is given by

$$
p(r_k, \theta_k) = \sum_{\theta_{k-1} \in \mathcal{M}} p(r_k, \theta_{0:k}) = \sum_{\theta_{k-1} \in \mathcal{M}} p(r_k \mid \theta_{0:k}) p(\theta_{0:k}).
$$

Fig. 1. Block diagram showing the plant $P$ subject to a deterministic input $u$, a noise process $v$, and a random fault signal $f(\theta)$. The shaded region is the fault detection scheme $(F, \delta)$, and the output $d$ is the decision or inference made by $(F, \delta)$. 
This density can be used to compute probabilities (3)–(6); for instance,
\[ P_k^{IP} = \sum_{\theta_{0:k}} P(\theta_{0:k}) \text{d} \theta_{0:k} \]
where \( k \in \mathcal{X} \), \( d_k \in \mathcal{D} \) and \( \theta_{0:k} \in \mathcal{M}^{k+1} \). Assume that each decision function \( \delta_k \) is a threshold function
\[ \delta_k(\theta_k) := I(|r_k| > \varepsilon_k), \quad \varepsilon_k > 0. \]
If \( z \) is some Gaussian random variable with \( z \sim \mathcal{N}(\mu, \sigma^2) \) and \( \varepsilon > 0 \), then \( \mathbb{P}(\varepsilon < z < \varepsilon) \) is given by
\[ \int^{\varepsilon}_{-\varepsilon} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(z-\mu)^2}{2\sigma^2}} \, dz. \]
where \( \mathbb{P} : \mathbb{R} \to [-1, 1] \) is the error function. Analytically, \( \mathbb{P}(\cdot) \) can be approximated by a rational function with maximum relative error less than \( 10^{-13} \) [14]. Evaluating the rational approximation requires a fixed number of floating point operations. Thus, the probability \( \mathbb{P}(\varepsilon < z < \varepsilon) \) can be computed accurately in \( \mathcal{O}(1) \) time. We conclude that once the densities \( p(\theta_k : k) \) and \( p(\theta_{0:k}) \) have been calculated (see Section III-B), the performance metrics can be evaluated with \( \mathcal{O}(k^4) \) calls to \( \mathbb{P}(\cdot) \) in \( \mathcal{O}(k^4) \) time.

B. Computational Procedure

Because of their linear structure, systems (1) and (2) can be combined into a single system
\[ \begin{align*}
\eta_{k+1} &= A_k \eta_k + B_{u,k} u_k + B_{v,k} v_k + B_{f,k} f_k (\theta_{0:k}), \\
\eta_k &= C_k \eta_k + D_{a,k} u_k + D_{v,k} v_k + D_{f,k} f_k (\theta_{0:k}),
\end{align*} \]
where \( \eta_k := (x_k, \xi_k) \) is the combined state. For each \( k \in \mathcal{X} \), define
\[ \delta_k := \mathbb{E}(\eta_k | \theta_{0:k} = \hat{\theta}_{0:k}) \text{ and } r_k := \mathbb{E}(r_k | \theta_{0:k} = \hat{\theta}_{0:k}), \]
Strictly speaking, we should write \( \hat{\eta}_k(\theta_{0:k}) \) and \( \hat{r}_k(\theta_{0:k}) \), but we omit this argument when the sequence \( \theta_{0:k} \) is clear from context. The sequences \{\( \hat{\eta}_k \)\} and \{\( \hat{r}_k \)\} are given by
\[ \begin{align*}
\hat{\eta}_{k+1} &= A_k \hat{\eta}_k + B_{u,k} u_k + B_{v,k} v_k + B_{f,k} f_k (\hat{\theta}_{0:k}), \\
\hat{\eta}_k &= C_k \hat{\eta}_k + D_{a,k} u_k + D_{v,k} v_k + D_{f,k} f_k (\hat{\theta}_{0:k}).
\end{align*} \]
Similarly, for each \( k \), define the conditional covariance matrices \( P_k := \text{var}(\eta_k | \theta_{0:k} = \hat{\theta}_{0:k}) \) and \( Q_k := \text{var}(r_k | \theta_{0:k} = \hat{\theta}_{0:k}). \)
Then, \{\( \hat{P}_k \)\} and \{\( \hat{Q}_k \)\} are given by
\[ \begin{align*}
P_{k+1} &= A_k \hat{P}_k A_k^* + B_{u,k} B_{u,k}^*, \\
Q_k &= C_k \hat{P}_k C_k^* + D_{a,k} D_{a,k}^*. \tag{13}
\end{align*} \]
Each update (from \( k \) to \( k+1 \)) of equations (12) and (13) takes constant time, so for a fixed final time \( T \in \mathcal{X} \) and a given mode sequence \( \hat{\theta}_{0:T} \), the sequences \( \hat{\eta}_{0:T} \) and \( \hat{Q}_{0:T} \) can be computed in \( \mathcal{O}(T) \) time. Since \( \{\theta_k \} \) is a Markov process,
\[ p(\theta_{0:k}) = p(\theta_k | \theta_{k-1}) p(\theta_{0:k-1}) = \Pi_{\theta_{k-1}} q(\theta_{0:k-1}), \]
and \( \mathbb{P}(\theta_{0:k} = \hat{\theta}_{0:k}) \) can be recursively computed from \( \mathbb{P}(\theta_{0:k-1} = \hat{\theta}_{0:k-1}) \) with a single multiplication. Since there are \( \mathcal{O}(T^3) \) mode sequences, the entire joint density (9) can be computed in \( \mathcal{O}(T^4+1) \) time.

C. Special Case

Recall that the Markov chain \( \{\theta_k \} \) is interpreted as the status of \( \ell \) components that fail randomly. We consider a special class of fault signals \( \{f_k(\cdot)\} \) of the form
\[ f_k(\theta_{0:k}) = \sum_{i=1}^{\ell} \lambda_i \left( k - T_i(\theta_{0:k}) \right), \quad k \in \mathcal{X}, \tag{14} \]
where each $\lambda_i$ is a deterministic function and $T_l(\theta_{0:k})$ is the time at which the $i$th component fails in the mode sequence $\theta_{0:k}$. If the $i$th component does not fail, we take $T_l(\theta_{0:k}) = \infty$. Assume that $\lambda_i(s) = 0$ for $s < 0$. In other words, $\lambda_i$ "switches on" when component $i$ fails. Since system (11) has linear structure, we can use superposition to significantly reduce the amount of computation needed to compute $\{\theta_t\}$ and $\{Q_k\}$.

For $i = 1, 2, \ldots, \ell$ and $\tau = 1, 2, \ldots, T$, define $\theta_{0:T}^{(i, \tau)}$ to be the mode sequence for which $T_i = \tau$ and $T_j = \infty$, for $i \neq j$. Suppose that for all such $(i, \tau)$ pairs, we set $\hat{\theta}_0 = 0$, $u_k = 0$, for all $k$, and simulate equation (12) with the input $\{f_{0:T}^{(i, \tau)}\}$ to obtain the corresponding conditional mean $\lambda_{0:T}^{(i, \tau)}$. Also, let $\hat{\rho}_{0:T}^{(0, 0)}$ be the result of simulating equation (12) with the original values of $\hat{\theta}_0$ and $u_0$ but with no faults. Then, the value of $\hat{\theta}_{0:T}$ corresponding to an arbitrary mode sequence $\theta_{0:T}$ can be obtained by superposing $\hat{\rho}_{0:T}^{(i, \tau)}$ for the appropriate pairs $(i, \tau)$. Using superposition the number of simulations needed is reduced from $O(T^{\ell+1})$ to $O(T^{\ell+1})$. Moreover, if the system (11) is LTV then for each $i$ and $j$, $\hat{\rho}_{0:T}^{(i, \tau)}$ can be shifted $\tau$ time-steps to obtain $\hat{\rho}_{0:T}^{(j, \tau + \tau)}$, which further reduces the number of simulations to $O(\ell)$. However, in these special cases, the error function must still be evaluated $O(T^{\ell+1})$ times to compute the performance metrics. These time-complexity results are summarized in Table I.

IV. APPLICATION: PITOT-STATIC PROBE

Nearly all aircraft use a pitot-static probe to determine airspeed $V$ and altitude $h$. Because these data are essential for flying, the pitot-static probe is integrated into the flight control feedback loop. These sensors are prone to a number of failures, such as icing and blockage, that cause them to produce incorrect values. If such a failure goes undetected, the autopilot system or the pilot may use the erroneous values to issue commands that cause the aircraft to crash. To avoid such disasters, large commercial aircraft, such as the Boeing 777 [2], [3], have multiple pitot-static probes in different locations. However, most aircraft designers have developed a set of standard operating procedures that allow safe recovery of the aircraft when a pitot-static probe failure is detected [15]. In this application we explore the detection of such faults by exploiting the analytical redundancy between airspeed, altitude, and flight path angle. Hansen et al. [16] present a similar example, which uses the methods of statistical change-point detection [17], [18] to model sensor faults.

A. System Description

Consider the fault detection problem shown in Fig. 2. The pitot tube measures the total pressure $p_t$, and the static port measures the static pressure $p_s$. These measurements are corrupted by Gaussian white noise processes, $v_i$ and $v_s$, and randomly occurring bias faults, $f_i$ and $f_s$. The airspeed and altitude are derived using the relations (see Fig. 3)

$$V = \phi(p_t, p_s) := \left[ \frac{\text{sign}(p_t - p_s)}{c_1 \left( \left( \frac{p_t - p_s}{p_0} \right)^4 + 1 \right)^{c_2} - 1} \right]^{1/2},$$

where the constants $c_1 = 44.331$ km, $c_2 = 0.1903$, $c_3 = 760.427$ m/s, $c_4 = 2/3$, and $p_0 = 101.325$ kPa model the troposphere ($h \leq 17$ km) [1]. We use the notation $V$ for the derived airspeed and $\dot{h}$ for the derived altitude to indicate that these quantities are corrupted by random disturbances and faults. Note $\phi$ actually gives the indicated airspeed, but we ignore this issue for the sake of simplicity.

The fault signals are defined as $f_i(t) := b_i I(t \geq T_i)$ and $f_s(t) := b_s I(t \geq T_s)$, for $t \geq 0$, where $b_i$ and $b_s$ are known, fixed biases and $T_i$ and $T_s$ are independent exponential random variables $T_i \sim \text{Exp}(\lambda_i)$ and $T_s \sim \text{Exp}(\lambda_s)$.

The dynamic portion of the fault detection scheme $F$ is contained in the shaded region of Fig. 2. The input $\gamma$ is the flight path angle of the aircraft, which we assume is measured exactly with no noises or faults. Consider the following analytical relationship between $V$, $h$, and $\gamma$:

$$h(t) = h(0) + \int_0^t V(\tau, \gamma(\tau)) d\tau = \int_0^t V(\tau) \sin(\gamma(\tau)) d\tau,$$

which is used to derive $\dot{h}$ from $\gamma$ and $\dot{V}$. The fault detection scheme attempts to detect the faults $f_i$ and $f_s$ by analyzing the difference $\dot{h} - \dot{h}$. However, as the noisy signal $\psi(\dot{V}, \gamma)$ passes through the integrator, the noise accumulates and $\dot{h}$ diverges from $\dot{h}$. To counteract this effect, a high-pass or "washout" filter of the form

$$W(s) = \frac{s}{s + a}, \quad a > 0,$$

is applied to the difference $\dot{h} - \dot{h}$ to produce the residual $r$. The drawback of using this filter is that it removes the DC component from the signal $\dot{h} - \dot{h}$. The decision function (not depicted in Fig. 2) is a threshold function with threshold $\varepsilon > 0$, and the same decision function is used at each $k \in \mathcal{K}$.

B. Applying the Framework

To apply the framework developed in Sections II and III, the plant $P$ must be LTV. As shown in Fig. 3, the map $\phi$ is only mildly nonlinear for modest changes in differential

<table>
<thead>
<tr>
<th>Fault Model</th>
<th>Simulations</th>
<th>Calls to erf(·)</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Markov Chain</td>
<td>$O(m^{\ell+1})$</td>
<td>$O(T^{m+1})$</td>
</tr>
<tr>
<td>Restricted Markov Chain</td>
<td>$O(T^\ell)$</td>
<td>$O(T^{\ell+1})$</td>
</tr>
<tr>
<td>Special case (LTV)</td>
<td>$O(T^\ell)$</td>
<td>$O(T^{\ell+1})$</td>
</tr>
<tr>
<td>Special case (LTI)</td>
<td>$O(\ell)$</td>
<td>$O(T^{\ell+1})$</td>
</tr>
</tbody>
</table>
The system (15) is discretized as follows: fix a sample time \( \Delta t \). For this analysis, we use the following parameter values: \( \Delta t = 0.05 \text{s} \), \( V = 45 \text{m/s} \), \( \gamma = 0.5^\circ \), \( h_0 = 200 \text{m} \), \( b_1 = -40 \text{Pa} \), \( b_y = 50 \text{Pa} \), and \( v_{1,k} \sim \mathcal{N}(0,0.297 \text{Pa}^2) \) and \( v_{y,k} \sim \mathcal{N}(0,0.297 \text{Pa}^2) \), for all \( k \in \mathcal{K} \). The component failure probabilities are \( q_i = q_s = 1.389 \cdot 10^{-7} \), which corresponds to a mean time-to-failure (MTTF) of 1000hrs [4].

To get a sense of how each fault affects system (15), we plot the change in the residual’s statistics due to a particular fault. Fig. 4 shows the change due to \( f_i \) occurring at \( T_f = 1 \text{ min} \), and Fig. 5 shows the residual due to \( f_s \) occurring at \( T_f = 1 \text{ min} \). Three cases are plotted: nominal (black), fault with \( a = 0.003 \) (blue), and fault with \( a = 0.0075 \) (green).

In each case, the region within three standard deviations of the mean is shaded in the appropriate color. Note that the green lines reach steady-state more quickly, but the blue lines achieve more separation between the nominal and faulty cases. Hence, there is a trade-off between how quickly the faults can be detected and how easily the fault is distinguished from the nominal case. The instantaneous jump shown in Fig. 5 is due to the direct feedthrough from \( f_s \) to \( r \) in (15). This makes it difficult to detect the fault between 3 and 8 min, when the residual changes sign.

In Fig. 6, the performance metrics \( \{P_k^{\text{FN}}\} \) (solid lines) and \( \{P_k^{\text{FP}}\} \) (dashed lines) are plotted for \( a = 0.003 \) and \( \varepsilon = 2 \text{m,3.5m,5m} \) (blue, green, and red, respectively). The curves \( \{P_k^{\text{FN}}\} \) and \( \{P_k^{\text{FP}}\} \) are omitted, because their values are approximately zero over the time window plotted. Hence, this plot does not provide information about the relative performance of each scheme when a fault occurs.

In Fig. 7, the performance metrics \( \{P_k^{\text{D}}\} \) (solid lines) and \( \{P_k^{\text{D}}\} \) (dashed lines) are plotted for \( a = 0.003 \) and \( \varepsilon = 2 \text{m,3.5m,5m} \) (blue, green, and red, respectively). Recall that \( P_k^{\text{D}} \) accounts for all possible faults up to time \( k \). For small \( k \), the instantaneous jump shown in Fig. 5 dominates, which causes a peak in the \( \{P_k^{\text{D}}\} \) curve at 4min. However, as \( k \) increases, these early faults begin to settle and there are more faults to consider, so the curve \( \{P_k^{\text{D}}\} \) becomes smoother. The values plotted in Fig. 6 and Fig. 7 allow us to directly compare the performance of different fault detection schemes and certify the time-varying reliability of the system under each scheme.

V. CONCLUSIONS

The reliability of safety-critical systems must be certified, but there is little work in the literature that rigorously analyzes the performance of model-based fault detection systems. The framework presented in this paper provides a class of model-based fault detection problems for which the performance can be computed analytically. It is shown that, under a reasonable set of assumptions, this computation can be carried out in polynomial time. This analysis is applied to a model of a pitot-static probe subject to randomly occurring bias faults with known magnitudes, which are detected using analytical redundancy. The data obtained from this analysis (shown in Fig. 6 and Fig. 7) can be used to certify the performance of a fault detection scheme.

Future work on this topic will include the study of more complex decision functions, such as the likelihood ratio test

\[ d_k = \mathcal{LRT}(r_{0,k}) := \mathbb{I} \left( \frac{p(r_{0,k} | H_{1,k})}{p(r_{0,k} | H_{0,k})} > \varepsilon_k \right), \]

and the up-down counter, which is defined by the recurrence

\[ c_{k+1} = c_k + C_{\text{up}} \mathbb{I}(r_k > \varepsilon_k) - C_{\text{down}} \mathbb{I}(r_k \leq \varepsilon_k), \]

where \( c_0 = 0, C_{\text{up}} \geq C_{\text{down}}>0 \), and \( \varepsilon_k, \tau_k > 0 \). The likelihood ratio test has desirable theoretical properties [12], and the up-down counter is commonly used in avionics applications [1].
Also, since the analysis in Section IV was carried out over a particular flight path, it would be useful to find the input $u$ that gives the worst fault detector performance. Similarly, if there is parametric model uncertainty in the plant $P$, it would be useful to find the worst-case uncertainty parameter value.

VI. ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. 0931931 entitled “CPS: Embedded Fault Detection for Low-Cost, Safety-Critical Systems” and by the National Aeronautics and Space Administration under Grant No. NNX07AC40A entitled “Reconfigurable Robust Gain-Scheduled Control for Air-Breathing Hypersonic Vehicles.” Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES