Probability Bounds for False Alarm Analysis of Fault Detection Systems

Bin Hu and Peter Seiler

Abstract— **Model-based fault detection methods have the potential to reduce the size, weight, and cost of safety-critical aerospace systems. One obstacle to the application of these methods is a lack of appropriate tools to efficiently certify their reliability. This paper focuses on the false alarm analysis of a general fault detection scheme. The main difficulty of the false alarm analysis is the time-correlations introduced by the plant dynamics and the fault detection filter. This paper proves product-type probability inequalities for general stationary zero-mean Gaussian processes. These inequalities are applied to provide converging bounds for the false alarm probability over a given time window. A numerical example is presented to demonstrate the proposed technique.**

I. Introduction

Safety-critical aerospace systems must be reliable and their reliability must also be certifiable. There is a need to certify the reliability of an aerospace safety-critical system with aviation authorities, e.g. the Federal Aviation Administration in the United States or the European Aviation Safety Agency. Commercial flight control electronics are required to have no more than 10^{-9} catastrophic failures per flight hour [2], [5]. The current design of safety-critical aerospace systems is based on physical redundancy [24], whose performance is relatively straightforward to certify using fault trees [15], [16]. However, physically redundant architectures increase the system size, weight, power, and cost. As a result, there have been efforts to develop analytical redundancy as an alternative approach to achieve fault tolerance, e.g. the oscillatory monitors on the Airbus A380 [12]. Model-based fault detection and isolation (FDI) is one method to realize analytical redundancy [4], [14], [6]. The recent AddSafe project in Europe [1] assessed the suitability of these advanced fault detection methods for commercial aircraft. In addition, model-based FDI could significantly improve the reliability of other safety-critical aerospace systems which cannot afford the size, weight and power associated with physical redundancy, e.g. Unmanned Aerial Vehicles and fly-by-wire in business/general aviation aircraft.

However, the certification of analytically redundant systems is challenging due to nonlinear, time-varying and uncertain aircraft dynamics, time-correlations in the residuals introduced by filtering, potentially complex decision functions, and stringent reliability requirements. One possible approach to validate the FDI performance involves linearizing the aircraft dynamics at many trim conditions in the flight envelope. Next, linear analysis tools are used to rigorously assess the FDI performance at these operating conditions. Finally, high fidelity Monte Carlo simulations can be performed to complement the linear analyses. This approach is similar to the existing procedure to validate the robustness and performance of flight control laws [18]. This paper seeks an analytical solution for the false alarm analysis of a FDI system at a trim condition. The main difficulty is the time-correlations introduced by plant dynamics and the FDI filter.

In the statistical process control (SPC) community, finite state Markov chain approximations are used to compute false alarm probabilities when the univariate control charting statistic is governed by a first-order autoregressive model [17]. This method could be useful for the false alarm analysis of a first-order FDI system. However, one drawback of the finite state Markov chain approximation is the "curse of dimensionality". This makes the Markov chain approximation computationally inefficient for more realistic FDI systems which are governed by higher-order dynamics.

Probability inequalities are another potential solution for false alarm analysis. Product-type inequalities were first derived by Sidak [23] to study confidence intervals of multivariate normal distributions. Glaz obtained a general form of the product-type inequalities based on the property of absolute value multivariate totally positive of order 2 (AMTP2) [8]. Product-type probability inequalities always provide better bounds than Bonferroni-type inequalities which are another class of probability bounds commonly used [13]. Although product-type probability inequalities have been used widely to study confidence intervals and time series models [10], they have not been used to study the performance of fault detection systems. The restriction of AMTP2 property is a strong condition for false alarm analysis. This paper derives a weaker condition where product-type inequalities can hold so that product-type inequalities can be used to analyze false alarm probabilities of general fault detection systems.

This paper formulates a false alarm analysis problem of a typical model-based fault detection system in Section II. Section III presents the main contribution of this paper: deriving product-type inequalities for general stationary zero-mean Gaussian processes. An efficient bounding method is also presented to solve the false alarm analysis problem. Finally, Section IV gives a numerical example to demonstrate the utility of the proposed method.

The analysis in this paper is complementary to Monte Carlo methods. In particular, Monte Carlo methods, such as the importance sampling and splitting techniques [21], have been used in rare event simulations. One potential drawback of the Monte Carlo method is the high computational power associated with the accuracy requirement and long length of

Bin Hu and Peter Seiler are with the Aerospace Engineering and Mechanics Department, University of Minnesota, Email: huxxx221@umn.edu; seiler@aem.umn.edu.

the time window. Both methods should be complementary to each other in order to provide a complete solution for the false alarm analysis problem.

II. Problem Formulation

Consider a typical fault detection scheme shown in Figure 1. G_θ denotes the monitored system, which depends on a parameter θ_k. $\theta_k \in \{0, 1\}$ denotes the status of the system at time k: $\theta_k = 0$ if G_θ is operational and $\theta_k = 1$ if a fault has occurred. G_θ is also subjected to known inputs u_k and unknown disturbances w_k. A fault detection scheme is used to monitor the status of G_θ. The fault detection scheme is compromised of two parts: a filter that generates a scalar residual carrying the information of the occurrence of the fault, and a decision function that generates a logic signal indicating the status of the monitored system. There are many approaches to design the FDI filter, e.g. observer based methods, parity equation based methods, parameter estimation methods, and advanced robust filter design methods [4], [14], [6]. In most cases, the residual r_k is generated based on noisy measurement y_k and known input u_k. Hence the residual r_k can be viewed as an output of the combined linear time-invariant (LTI) system comprised of the monitored system and the FDI filter.

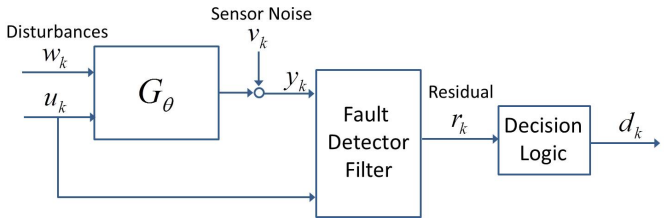

Fig. 1. Block diagram of a typical fault detection scheme.

The residual generator is typically designed in a way such that r_k is small when $\theta_k = 0$, and large when a fault occurs. Based on r_k, the decision logic generates a signal d_k to indicate the status of G_θ, i.e. $d_k = 1$ if a fault has been detected and $d_k = 0$ otherwise. There are many approaches for designing decision function logic, such as thresholding, statistical testing methods, and fuzzy logic [14], [6]. Our analysis focuses on constant thresholding:

$$d_k := \begin{cases} 0 & \text{if } |r_k| \le T \\ 1 & \text{else} \end{cases} \tag{1}$$

A fault is declared when r_k exceeds the threshold T. Thresholding is widely used in industrial applications due to its simplicity. The restriction to constant thresholds can also be viewed as a steady-state approximation for time-varying thresholds. This paper focuses on the difficulty of analysis introduced by the complicated state-space model that governs r_k. Therefore, a simple decision function is considered. This case provides a foundation for further research on FDI schemes with more complex decision functions.

For safety-critical aerospace systems, system reliability requirements are typically specified over a specified time interval, e.g. flight control systems certified with the FAA are required to have less than 10^{-9} catastrophic failures per flight hour [5]. These system level requirements indicate that the false alarm probability which is of interest in this paper should also be specified over a time interval. This motivates the following definition of false alarm probability:

Definition 1: The N-step false alarm probability, denoted P_N, is the conditional probability that $d_k = 1$ for some k in $1 \le k \le N$, given that $\theta_k = 0$ for all k in $1 \le k \le N$.

A false alarm is an event where the FDI scheme declares a fault when no fault has occurred within the N-step window. To compute P_N, a specific mathematical model for r_k is needed. False alarm analysis only considers the fault free case where $\theta_k = 0$. As mentioned in the introduction this paper seeks an analytical solution for the false alarm analysis of a FDI system at a trim condition. Hence finally, assume G_θ is in the trim condition so that $u_k = 0$ and the system is in the steady state. Then r_k is governed by the following discrete-time state-space model:

$$\begin{aligned} x_{k+1} &= Ax_k + Bn_k \\ r_k &= Cx_k + Dn_k \end{aligned} \tag{2}$$

Here, $x_k \in \mathbb{R}^h$, $n_k \in \mathbb{R}^l$, $A \in R^{h\times h}$, $B \in \mathbb{R}^{h\times l}$, $C \in \mathbb{R}^{1\times h}$, and $D \in \mathbb{R}^{1\times l}$. n_k is an independent and identically distributed (IID) Gaussian process with $n_k \sim \mathcal{N}(0, \Sigma)$. This model is quite general since it is valid for the following two classes of problems:

1) The unknown disturbances w_k are deterministic and the FDI filter realizes disturbances decoupling so that w_k does not affect r_k.
2) The unknown disturbances w_k are stochastic IID Gaussian process.

For the first case define $n_k = v_k$. For the second case define $n_k = \left[w_k; v_k\right]$ so n_k contains the combined effects of both the sensor noise and the stochastic disturbances. Then for both cases Equation 2 is a suitable model describing the dynamics governing r_k. Assume the system is stable, hence all eigenvalues of A have magnitude strictly less than 1. The steady-state assumption implies that for any k, the marginal distribution of $\{r_i, r_{i+1}, \ldots, r_{i+k}\}$ does not depend on i. Since n_k is IID Gaussian noise with zero mean, r_k is a strictly stationary zero-mean Gaussian process. The steady state covariance matrix Σ_x of the random vector x_k can be solved from the Lyapunov equation:

$$\Sigma_x = A\Sigma_x A^T + B\Sigma B^T \tag{3}$$

Define $\mathbf{R}_N = \begin{bmatrix} r_1 & r_2 & r_3 & \cdots & r_N \end{bmatrix}^T$. $\mathbf{R}_N$ is a Gaussian random vector. Let Λ_N denote the covariance matrix of $\mathbf{R}_N$. r_k is a strictly stationary Gaussian process with zero mean hence the covariance matrix of $\mathbf{R}_N$ has a Toeplitz structure:

$$\Lambda_N(i,k) = \begin{cases} C\Sigma_x C^T + D\Sigma D^T & \text{if } i = k \\ CA^{|i-k|}\Sigma_x C^T + CA^{|i-k|-1}B\Sigma D^T & \text{else} \end{cases} \tag{4}$$

Therefore the probability density function of the residual

vector $\mathbf{R}_N$ has the form:

$$f_N(\mathbf{R}_N) = \frac{1}{\sqrt{(2\pi)^N |\Lambda_N|}} e^{-\frac{1}{2}\mathbf{R}_N^T \Lambda_N^{-1} \mathbf{R}_N} \tag{5}$$

To simplify notation, $f_N(r_1, r_2, \cdots, r_N) := f_N(\mathbf{R}_N)$, so there is no confusion when writing an expression such as $f_N(c, r_2, \cdots, r_N)$. The lower index N emphasizes the dimension of the residual vector $\mathbf{R}_N$.

It is straightforward to compute the probability of an intersection of events based on a probability integral. Thus, based on Definition 1, it is helpful to express P_N as follows:

$$\begin{aligned} P_N &= P[\cup_{k=1}^N \{|r_k| > T\}] \\ &= 1 - P[\cap_{k=1}^N \{|r_k| \le T\}] \\ &= 1 - \int_{-T}^{T} f_N(\mathbf{R}_N) d\mathbf{R}_N \end{aligned} \tag{6}$$

To find P_N from Equation 6, three main difficulties need to be handled. First, r_k are time-correlated due to the dynamics of Equation 2. Second, N could be relatively large for our analysis. For example, a system with a 100Hz sample rate has $N = 3.6 \times 10^5$ samples per hour. Finally, for good FDI performance, safety-critical systems require P_N to be extremely small, and hence the accuracy requirement is very stringent for false alarm analysis. Therefore, generally, P_N can not be solved by directly numerical integration from Equation 6. Importance sampling is also time consuming due to the large value of N. As a complementary approach, Section III derives new probability bounds for stationary zero-mean Gaussian processes in order to solve for P_N.

III. False Alarm Analysis

This section presents a theoretical method to analyze the false alarm probability, P_N. Section III-A proves that product-type inequalities hold for general stationary zero-mean Gaussian processes and applies this result to obtain bounds for false alarm probability, P_N. Section III-B explains how the probability bounds can provide accurate estimates of false alarm probability based on convergence and perturbation arguments. Related work in the literature is discussed at the end of each subsection.

A. Probability Bounds for False Alarm Probability

For simplicity, denote $Q_N := 1 - P_N$. Q_N is the conditional probability that no alarm is declared within the N-step window given there is no fault. From Equation 6, it is clear that $Q_N = \int_{-T}^{T} f_N(\mathbf{R}_N) d\mathbf{R}_N$. It is beneficial to provide tight upper bounds for the false alarm probability, P_N. Hence, we seek lower bounds for Q_N. The next lemma is used to derive a lower bound on Q_N.

Lemma 1: Suppose r_k is a stationary zero-mean Gaussian process. Then:

$$\begin{aligned} &P[|r_N| \le T \mid |r_{N-1}| \le T, \cdots, |r_2| \le T, |r_1| \le c] \ge \\ &P[|r_N| \le T \mid |r_{N-1}| \le T, \cdots, |r_2| \le T] \end{aligned} \tag{7}$$

where $N \ge 3$ and $c \ge 0$.

Proof: The right side of Inequality 7 can be expressed as $\frac{Q_{N-1}}{Q_{N-2}}$ by the stationarity of r_k. Define $F_N(c) = P[|r_N| \le T, |r_{N-1}| \le T, \cdots, |r_2| \le T, |r_1| \le c]$. The left side of Inequality 7 can be expressed as $\frac{F_N(c)}{F_{N-1}(c)}$. Hence it is sufficient to show that $G(c) = F_N(c)Q_{N-2} - F_{N-1}(c)Q_{N-1} \ge 0$.

To simplify notation, define $H(c) : [0, +\infty) \to [0, 1]$ as follows:

$$H(c) = P[|r_N| \le T \mid |r_{N-1}| \le T, \cdots, |r_2| \le T, r_1 = c] \tag{8}$$

where $N \ge 3$. $H(c)$ has an analytical expression:

$$H(c) = \frac{\int_{-T}^{T} \cdots \int_{-T}^{T} f_N(c, r_2, \cdots, r_N) dr_2 \cdots dr_N}{\int_{-T}^{T} \cdots \int_{-T}^{T} f_{N-1}(c, r_2, \cdots, r_{N-1}) dr_2 \cdots dr_{N-1}} \tag{9}$$

For simplicity, denote the numerator and the denominator as $U(c)$ and $D(c)$, respectively.

To show $G(c) \ge 0$, first notice $\lim_{c \to +\infty} F_N(c) = Q_{N-1}$ and $F_N(0) = 0$, so $\lim_{c \to +\infty} G(c) = 0$ and $G(0) = 0$. Apply the fundamental theorem of calculus (Theorem 6.19 in [22]) to compute the derivative of $G(c)$ with respect to c as follows:

$$\frac{dG}{dc} = 2D(c)Q_{N-2}\left(H(c) - \frac{Q_{N-1}}{Q_{N-2}}\right) \tag{10}$$

By Lemma 3, which is stated and proved in the appendix, $H(c)$ is a non-increasing function. For fixed N, $\frac{Q_{N-1}}{Q_{N-2}}$ is a constant. Hence only two cases need to be considered. The first case is that there exists a certain constant c^* such that $\left(H(c) - \frac{Q_{N-1}}{Q_{N-2}}\right)$ is nonnegative for $c \le c^*$ and nonpositive for $c > c^*$. Since $D(c)Q_{N-2} \ge 0$, $\lim_{c \to +\infty} G(c) = 0$ and $G(0) = 0$, the only possible result is that $G(c)$ first increases from 0 as c increases from 0 to c^*, and then decreases to 0 as c approaches $+\infty$. So $G(c) \ge 0$ for all $c \ge 0$. The second case is that $\left(H(c) - \frac{Q_{N-1}}{Q_{N-2}}\right)$ is nonpositive for all c. Since $D(c)Q_{N-2} \ge 0$, $\lim_{c \to +\infty} G(c) = 0$ and $G(0) = 0$, the only possible result is that $G(c) = 0$ for all c. For both cases, $G(c) \ge 0$ for all $c \ge 0$. This completes the proof. ■

The main result can now be stated to provide upper bounds for false alarm probability P_N.

Theorem 1: Suppose r_k is a stationary zero-mean Gaussian process. Given N define

$$\gamma_N^{(k)} = \left(\frac{Q_k}{Q_{k-1}}\right)^{N-k} Q_k \tag{11}$$

for $2 \le k \le N$. Then

$$P_N \le 1 - \gamma_N^{(k)} \tag{12}$$

and the bounds $\gamma_N^{(k)}$ are non-decreasing in k.

Proof: It is trivial to verify $\gamma_N^{(N)} = Q_N$. Now set $c = T$ in Lemma 1 to get:

$$\frac{Q_N}{Q_{N-1}} \ge \frac{Q_{N-1}}{Q_{N-2}} \tag{13}$$

where $N \geq 3$. The monotonicity of $\gamma_N^{(k)}$ can be shown as follows:

$$\gamma_N^{(k)} = \left(\frac{Q_k}{Q_{k-1}}\right)^{N-k} Q_k \leq \left(\frac{Q_{k+1}}{Q_k}\right)^{N-k} Q_k = \gamma_N^{(k+1)} \tag{14}$$

Hence $\gamma_N^{(1)} \leq \ldots \leq \gamma_N^{(N)} = Q_N$, i.e. $\gamma_N^{(k)}$ is a lower bound on Q_N for $k = 2, \ldots, N$. The upper bounds on P_N follow immediately from $P_N = 1 - Q_N$. ■

This result provides a sequence of monotonically converging product-type bounds for the false alarm probability P_N. Terms involving $\gamma_N^{(k)}$ for small values of k can be used to bound P_N for large N. $\gamma_N^{(k)}$ is computed based on Q_k and Q_{k-1}. For large N computing bounds on P_N using this result is significantly more efficient than directly computing P_N. For example, Q_1 is a one-dimensional Gaussian integral, and it can be accurately computed from the error function, e.g. `erf` in `Matlab`. Q_2 and Q_3 correspond to two and three dimensional Gaussian integrals, respectively. These integrals can also be efficiently computed to within machine (double) precision using the `Matlab` function `mvncdf` [7]. Hence $\gamma_N^{(3)}$ can be computed trivially to bound P_N.

Glaz proved a similar product-type inequality assuming the random variable satisfies the technical AMTP2 condition [8], [10]. Theorem 1 generalizes this result and demonstrates that the product-type inequality holds for general stationary zero-mean Gaussian processes without assuming the AMTP2 condition which is strong and not easy to check as N increases. Product-type inequalities have been applied to obtain upper bounds for confidence intervals on multivariate normal variables [10]. However, in those studies, N is always chosen to be smaller than 100 and P_N is not an extremely small number close to 0. False alarm analysis faces a significantly different situation where N is much larger and P_N is much smaller with a higher computational accuracy requirement. Hence, the application of these inequalities in false alarm analysis should be carefully justified. It must be clarified when these probability inequalities can provide accurate estimates for false alarm probability.

B. Convergence of the Probability Bounds

As shown in the proof of Theorem 1, $\gamma_N^{(N)} = Q_N$. Hence for fixed N the bounds $\gamma_N^{(k)}$ are non-decreasing in k and convergent to Q_N. To see the convergence rate, the following result is needed:

Corollary 1: Suppose r_k is a stationary zero-mean Gaussian process, then $\frac{Q_{N+1}}{Q_N}$ is a monotonically non-decreasing, convergent sequence as a function of N.

Proof: The monotonicity of $\frac{Q_{N+1}}{Q_N}$ was demonstrated in the proof of Theorem 1. Moreover, $\frac{Q_{N+1}}{Q_N} \leq 1$ follows since $\cap_{i=1}^{N+1}\{|r_i| \leq T\} \subset \cap_{i=1}^{N}\{|r_i| \leq T\}$. It is also true that $\frac{Q_{N+1}}{Q_N}$ is nonnegative. A bounded monotone sequence in $\mathbb{R}$ converges to a finite limit (Theorem 3.14 in [22]). Therefore, $\lim_{N\to\infty} \frac{Q_{N+1}}{Q_N}$ exists and is an upper bound for any element of the sequence. ■

This corollary states that for a stationary zero-mean Gaussian process, Q_N monotonically converges to a geometric series. We can compute Q_N for relatively small N. As long as $\frac{Q_{N+1}}{Q_N}$ converges to a constant, Q_N can be treated as a geometric series. Based on the geometric ratio, computing Q_N for large N is simple. For example, suppose Q_N becomes a geometric series for $N \geq k$. Then Q_N can be computed as $Q_N = Q_k(\frac{Q_{k+1}}{Q_k})^{N-k} = \gamma_N^{(k+1)}$. Hence, in this case, $\gamma_N^{(k+1)}$ can be used as an accurate estimate of Q_N instead of just providing a bound.

If the convergence rate to the geometric series is extremely slow, this geometric series approximation idea may not be accurate. However, for false alarm analysis, Q_N typically has fast convergence to a geometric series. To see this heuristically, suppose the threshold for the decision logic is $T = \infty$. Then, $Q_N = 1$ for any N. Q_N is a geometric series from the first term. For fault detection problems, T is chosen to be relatively large in order to make the false alarm probability small. In this case, Q_N is a minor perturbation from those where $T = \infty$. It is expected that Q_N converges to a geometric series very quickly. Hence, for a wide class of problems where the convergence rate is fast enough, we can estimate Q_N based on Q_1, Q_2 and Q_3, all of which are easily computable. Theorem 1 turns the previous argument into a precise mathematical result that gives upper bounds on false alarm probability P_N.

Glaz has obtained a result similar to Corollary 1 when r_k is a moving sum of IID Gaussian random variables. He applied $\gamma_N^{(k)}$ as approximations of Q_N when product-type inequalities can not be proved [9]. Here, Corollary 1 holds as long as r_k is a stationary zero-mean Gaussian process regardless of whether the dynamic model governing r_k is a moving average model or not.

IV. Numerical Examples

This section presents numerical examples to demonstrate the results. All examples are based on a per-hour false alarm requirement for a system whose sample rate is $100Hz$. Thus there are $N = 3.6 \times 10^5$ sample frames per hour. In addition, a steady state condition is assumed for false alarm analysis.

A. Benchmark Problem: First Order Process

First, the probability bounding technique is applied to a benchmark problem which can also be solved by finite state Markov chain approximation. For this case the FDI scheme is assumed to be a first order autoregressive model:

$$r_{k+1} = ar_k + n_k \tag{15}$$

where $0 \leq a < 1$. n_k is an IID Gaussian process with $n_k \sim \mathcal{N}(0, 1)$. For a first-order process, the finite-state Markov chain approach can be used to efficiently compute the false alarm probability P_N. A basic review of this technique is included in the appendix. As comparisons, $\gamma_N^{(2)}$ and $\gamma_N^{(3)}$ are also used to compute estimates of P_N. The one-step false alarm probability P_1 is commonly referred to as the

false alarm rate (FAR), which can be expressed as a one-dimensional Gaussian integral:

$$P_1 = 1 - \int_{-T\sqrt{1-a^2}}^{T\sqrt{1-a^2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \tag{16}$$

Table I shows the hourly false alarm probabilities P_N computed from Markov chain approximation for several values of (a, T) all chosen to have $P_1 = 10^{-11}$. The values of (a, T) that give $P_1 = 10^{-11}$ are those that satisfy $T\sqrt{1-a^2} = 6.807$. Table I also shows the first two bounds $\gamma_N^{(2)}$ and $\gamma_N^{(3)}$ given by Theorem 1. These bounds use Q_1 computed by the `erf` function as well as Q_2 and Q_3 computed by the `mvncdf` function. Table I shows that, for a fixed one-step FAR P_1, an increase in a leads to a decrease in the hourly probability P_N. In other words, increased correlation leads to a decrease in the per hour false alarm probability. Note that $1 - Q_1^N \approx NP_1 = 3.6 \times 10^{-6}$ is an estimate of P_N obtained by assuming independent residuals, i.e. ignoring the time correlations. Based on Table I, the time correlation has negligible effects ($P_N \approx NP_1$) for $a \leq 0.8$ and the effects of correlation appear for larger values of a. For $a = 0.999$, P_N is only 4% of the value NP_1. The bounds $1-\gamma_N^{(i)}$ are an improvement on the simple uncorrelated value NP_1. It is notable that the bounds $1-\gamma_N^{(i)}$ also become less accurate for values of a near 1. The first bound $1 - \gamma_N^{(2)}$ is very accurate for $a < 0.9$ and of reasonable accuracy up to $a < 0.99$. There is not much difference between the first two bounds. This supports the perturbation argument made in Section III-B.

a	T	P_N	$1 - \gamma_N^{(3)}$	$1 - \gamma_N^{(2)}$
0	6.81	3.60×10^{-6}	3.60×10^{-6}	3.60×10^{-6}
0.7	9.53	3.59×10^{-6}	3.59×10^{-6}	3.60×10^{-6}
0.8	11.3	3.52×10^{-6}	3.52×10^{-6}	3.53×10^{-6}
0.9	15.6	3.17×10^{-6}	3.17×10^{-6}	3.20×10^{-6}
0.99	48.3	9.64×10^{-7}	1.18×10^{-6}	1.36×10^{-6}
0.999	152	1.40×10^{-7}	3.40×10^{-7}	4.45×10^{-7}

TABLE I

FALSE ALARM PROBABILITIES AND BOUNDS FOR A FIRST ORDER PROCESS. $N = 3.6 \times 10^5$, $P_1 = 10^{-11}$ FOR EACH CHOICE OF (a, T).

B. Higher order systems

This section presents an example of a second order process which can not be solved by the Markov chain approximation. Based on Table I, it is reasonable to hypothesize that the performance of the bounds is dominated by the spectral radius of the system state matrix. The example in this section is used to confirm this hypothesis. The FDI residual is assumed to be a second order process:

$$r_{k+1} = \phi_1 r_k + \phi_2 r_{k-1} + n_k \tag{17}$$

where ϕ_1 and ϕ_2 satisfy the stability conditions: $\phi_1 + \phi_2 < 1$, $\phi_2 - \phi_1 < 1$ and $-1 < \phi_2 < 1$. These conditions ensure that the system poles all have magnitude strictly less than 1. n_k is an IID Gaussian process with $n_k \sim \mathcal{N}(0, 1)$.

Table II shows the bounds $\gamma_N^{(2)}$ and $\gamma_N^{(3)}$ for the hourly false alarm probabilities P_N for real poles z_1 and z_2 with corresponding threshold T chosen to satisfy FAR $P_1 = 10^{-11}$. z_1 is the dominant pole. Hence $\phi_1 = z_1 + z_2$ and $\phi_2 = -z_1 z_2$. The process described by Equation 17 can be recast into a state-space model governed by Equation 2 with $A = \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 \end{bmatrix}^T$, $C = \begin{bmatrix} 1 & 0 \end{bmatrix}$ and $D = 0$. Then from Equation 3 and 4, the steady state variance of r_k can be computed as $\Lambda_1 = C\Sigma_x C^T$. $T = 6.807\sqrt{\Lambda_1}$ so that $P_1 = 10^{-11}$. Again $1 - Q_1^N \approx NP_1 = 3.6 \times 10^{-6}$ is an estimate of P_N obtained by assuming independent residuals, i.e. ignoring the time correlations. Compare Tables I and II, it is shown that the second order process can be approximated by a first order process with a single pole z_1 if z_2 is significantly smaller than z_1. When z_2 approaches to z_1, the effect of correlation becomes significant and the bounds decrease. The last line of Table II supports the hypothesis that larger spectral radius of the system state matrix leads to poorer performance of the bounds since the two bounds become significantly different for $z_1 = 0.999$ and the convergence argument is not convincing for this case.

z_1	z_2	T	$1 - \gamma_N^{(3)}$	$1 - \gamma_N^{(2)}$
0.7	0.1	10.3	3.56×10^{-6}	3.56×10^{-6}
0.7	0.7	22.8	2.79×10^{-6}	2.80×10^{-6}
0.99	0.1	53.6	1.09×10^{-6}	1.23×10^{-6}
0.99	0.99	3412	9.97×10^{-8}	9.99×10^{-8}
0.999	0.1	169	3.18×10^{-7}	4.01×10^{-7}

TABLE II

FALSE ALARM BOUNDS FOR A SECOND ORDER PROCESS. $N = 3.6 \times 10^5$ AND $P_1 = 10^{-11}$ FOR EACH CHOICE OF (z_1, z_2, T).

Table III presents the bounds $\gamma_N^{(2)}$ and $\gamma_N^{(3)}$ for complex poles $ae^{\pm j\psi}$ with corresponding threshold T chosen to satisfy FAR $P_1 = 10^{-11}$. And $\phi_1 = 2a\cos\psi$ and $\phi_2 = -a^2$.

a	ψ	T	$1 - \gamma_N^{(3)}$	$1 - \gamma_N^{(2)}$
0	0	6.81	3.60×10^{-6}	3.60×10^{-6}
0.7	0	22.8	2.79×10^{-6}	2.80×10^{-6}
0.7	$\frac{\pi}{4}$	10.4	3.58×10^{-6}	3.58×10^{-6}
0.7	$\frac{\pi}{2}$	7.81	3.59×10^{-6}	3.59×10^{-6}
0.7	$\frac{3\pi}{4}$	10.4	3.58×10^{-6}	3.58×10^{-6}
0.7	π	22.8	2.79×10^{-6}	2.80×10^{-6}
0.99	0	3412	9.97×10^{-8}	9.99×10^{-8}
0.99	$\frac{\pi}{4}$	48.5	3.57×10^{-6}	3.57×10^{-6}
0.99	$\frac{\pi}{2}$	34.4	1.84×10^{-6}	3.59×10^{-6}
0.99	$\frac{3\pi}{4}$	48.5	3.57×10^{-6}	3.57×10^{-6}
0.99	π	3412	9.97×10^{-8}	9.99×10^{-8}

TABLE III

FALSE ALARM BOUNDS FOR A SECOND ORDER PROCESS. $N = 3.6 \times 10^5$ AND $P_1 = 10^{-11}$ FOR EACH CHOICE OF (a, ψ, T).

The results for $\psi = 0$ again shows that repeated poles significantly increase the effect of correlation and decrease P_N. Tables III shows the bounds are symmetric about $\psi = \frac{\pi}{2}$. For ψ near $\frac{\pi}{4}$ or $\frac{3\pi}{4}$, the time correlation has negligible effects ($P_N \approx NP_1$). For ψ near 0 or π, the system can be approximated by a second order system with repeated real poles. Compare Table I and III, we can see that, as ψ approaches to $\frac{\pi}{2}$, the second order system starts to behave like a first order system with single pole a. It is interesting to notice that for this case, the bound $1-\gamma_N^{(3)}$ works well to provide estimates of P_N but the bound $1-\gamma_N^{(2)}$ fails to work. Finally it is worth mentioning that one can check numerically that the first order autocorrelation $\frac{E[r_k r_{k+1}]}{E[r_k^2]}$ is closed to 1 for all cases where time correlations cannot be ignored.

One thing worth noting is that the second order process is AMTP2 only when $\phi_2 < 0$. It is obvious that when both poles are real and positive, the process is not AMTP2. The results in Section III-A demonstrate that the product-type inequalities in [8] are still valid even if the AMTP2 conditions fail to be satisfied.

V. Conclusion

This paper analyzed the false alarm probability over a given time window for a general fault detection system. Product-type probability inequalities are proved for stationary zero-mean Gaussian processes. These inequalities are applied to provide computationally cheap bounds on the false alarm probability. A heuristic argument is made to demonstrate the conditions under which the bounds actually provide an accurate estimate of the false alarm probability. Numerical examples are presented to demonstrate that the proposed method can provide accurate results for false alarm probability for a wide class of FDI systems. The effects of correlation and pole positions have been investigated. Various applications of the product-type inequalities proved in this paper are being studied.

VI. Acknowledgments

This work was partially supported by the National Science Foundation under Grant No. NSF-CMMI-1254129 entitled "CAREER: Probabilistic Tools for High Reliability Monitoring and Control of Wind Farms". This work was also partially supported by NASA under Grant No. NRA NNX12AM55A entitled "Analytical Validation Tools for Safety Critical Systems Under Loss-of-Control Conditions", Dr. Christine Belcastro technical monitor. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the NSF or NASA.

References

[1] ADDSAFE: Advanced fault diagnosis for sustainable flight guidance and control. http://addsafe.deimos-space.com/, 2012. European 7th Framework Program.

[2] R.J. Bleeg. Commercial jet transport fly-by-wire architecture considerations. In *AIAA/IEEE Digital Avionics Systems Conference*, pages 399–406, 1988.

[3] D. Brook and D. A. Evans. An approach to the probability distribution of CUSUM run length. *Biometrika*, 59(3):539–549, 1972.

[4] J. Chen and R.J. Patton. *Robust Model-Based Fault Diagnosis for Dynamic Systems*. Kluwer, 1999.

[5] R.P.G. Collinson. *Introduction to Avionic Systems*. Kluwer, 2003.

[6] S.X. Ding. *Model-Based Fault Diagnosis Techniques: Design Schemes, Algorithms, and Tools*. Springer-Verlag, 2008.

[7] A. Genz. Numerical computation of rectangular bivariate and trivariate normal and t probabilities. *Statistics and Computing*, 14:251–260, 2004.

[8] J. Glaz and B. Johnson. Probability inequalities for multivariate distributions with dependence structures. *Journal of the American Statistical Association*, 79(386):436–440, 1984.

[9] J. Glaz and B. Johnson. Boundary crossing for moving sums. *Journal of Applied Probability*, 25(1):pp. 81–88, 1988.

[10] J. Glaz and N. Ravishanker. Simultaneous prediction intervals for multiple forecasts based on Bonferroni and product-type inequalities. *Statistics & Probability Letters*, 12(1):57–63, 1991.

[11] G.H. Golub and C.F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, 3rd edition, 1996.

[12] P. Goupil. Oscillatory failure case detection in the A380 electrical flight control system by analytical redundancy. *Control Engineering Practice*, 18(9):1110–1119, 2010.

[13] D. Hoover. Comparisons of improved Bonferroni and Sidak/Slepian bounds with applications to normal markov processes. *Communications in Statistics - Theory and Methods*, 19(5):1623–1637, 1990.

[14] R. Isermann. *Fault-Diagnosis Systems: An Introduction from Fault Detection to Fault Tolerance*. Springer-Verlag, 2006.

[15] M. Krasich. Use of fault tree analysis for evaluation of system-reliability improvements in design phase. *IEEE Proc. Reliability and Maintainability Symposium*, pages 1–7, 2000.

[16] W.S. Lee, D.L. Grosh, A.F. Tillman, and C.H. Lie. Fault tree analysis, methods, and applications: a review. *IEEE Trans. on Reliability*, 34(3):194–203, 1985.

[17] J. M. Lucas and M. S. Saccucci. Exponentially weighted moving average control schemes: properties and enhancements. *Technometrics*, 32(1):pp. 1–12, 1990.

[18] J. Renfrow, S. Liebler, and J. Denham. F-14 flight control law design, verification, and validation using computer aided engineering tools. In *IEEE Conference on Control Applications*, pages 359–364, 1994.

[19] S. W. Roberts. Control chart tests based on geometric moving averages. *Technometrics*, 1(3):pp. 239–250, 1959.

[20] H.L. Royden. *Real Analysis*. Macmillan, 3rd edition, 1988.

[21] G. Rubino and B. Tuffin. *Rare Event Simulation using Monte Carlo Methods*. Wiley, 2009.

[22] W. Rudin. *Principles of Mathematical Analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill Education, 1953.

[23] Z. Sidak. On multivariate normal probabilities of rectangles: Their dependence on correlations. *The Annals of Mathematical Statistics*, 39(5):pp. 1425–1434, 1968.

[24] Y.C. Yeh. Triple-triple redundant 777 primary flight computer. In *Proceedings of the 1996 IEEE Aerospace Applications Conference*, pages 293–307, 1996.

Appendix

A. Supporting Lemmas

The next two lemmas are used to prove Lemma 1 in Section III-A.

Lemma 2: Suppose $\Lambda_N \in \mathbb{R}^{N\times N}$ is a symmetric positive definite Toeplitz matrix where $\Lambda_N(i,1) = \beta_{i-1}$. Then:

$$\Lambda_N^{-1}(1,1) - \Lambda_{N-1}^{-1}(1,1) \geq 0 \tag{18}$$

where $\Lambda_N^{-1}(i,j)$ denotes the (i,j)-th entry of matrix Λ_N^{-1}.

The proof is based on a well-known result regarding the inverse of a symmetric positive definite Toeplitz matrix (Section $4.7.4$ in [11]). It is a straightforward proof and hence the detailed proof is omitted here. Lemma 2 leads to the following fact stated in Lemma 3:

Lemma 3: Suppose r_k is a stationary zero-mean Gaussian process, and use the notation $H(c)$, $U(c)$ and $D(c)$ introduced in the proof of Lemma 1 in Section III-A. Then, $H(c)$ is a non-increasing function of c.

Proof: The first order derivative of $H(c)$ with respect to c is computed as follows:

$$\frac{dH}{dc} = \frac{\frac{dU}{dc}D(c) - U(c)\frac{dD}{dc}}{D^2(c)} \tag{19}$$

Equation 19 can be further simplified based the Lebesgue's dominated convergence theorem (Section 4.4 in [20]) and the final result is:

$$\frac{dH}{dc} = \frac{-cU(c)}{D(c)}\left(\Lambda_N^{-1}(1,1) - \Lambda_{N-1}^{-1}(1,1)\right) \tag{20}$$

Since Λ_N is a symmetric positive definite Toeplitz matrix, by Lemma 2, we have $\Lambda_N^{-1}(1,1) - \Lambda_{N-1}^{-1}(1,1) \geq 0$. Since $U(c)$, $D(c)$ and c are all nonnegative, Equation 20 shows $\frac{dH}{dc} \leq 0$. Therefore, $H(c)$ is a non-increasing function. ■

B. Review of Finite State Markov Chain Approximation

The first-order autoregressive process in Equation 15 is closely related to exponentially weighted moving average (EWMA) charts in the statistical process control (SPC) community going back to the initial work by Roberts [19]. The finite state Markov chain approximation have been used to analyze the performances of EWMA charts [3], [17]. The idea is based on the fact that Equation 15 ensures r_k to be a Markov chain. For finite state Markov chains, the false alarm probability P_N has an approximately geometric distribution and the related distribution parameter has an eigenvalue interpretation. This is a consequence of the Perron-Frobenius theorem. The key result in [3] can be restated in a slightly different way as the following theorem:

Theorem 2: Let $M \in \mathbb{R}^{m\times m}$ be a finite substochastic matrix which is irreducible and aperiodic. Let $\mathbf{1} \in \mathbb{R}^m$ be a vector whose elements are all 1. Let $\pi_0 \in \mathbb{R}^m$ be a vector whose elements are nonnegative. Let $\lambda_1 \geq \lambda_2$ be the eigenvalues of M with greatest modulus. Y_L and Y_R are the corresponding eigenvectors $Y_L^T M = \lambda_1 Y_L^T$, $MY_R = \lambda_1 Y_R$. Then λ_1 is real and has multiplicity 1. $\lambda_1 = 1$ if M is stochastic and $0 < \lambda_1 < 1$ otherwise. Moreover, the following approximation holds as $N \to \infty$:

$$\pi_0^T M^{N-1}\mathbf{1} = c_1\lambda_1^{N-1}\left[1 + \mathcal{O}\left(\left(\frac{|\lambda_2|}{|\lambda_1|}\right)^{N-1}\right)\right] \tag{21}$$

where $c_1 = \frac{\pi_0^T Y_R Y_L^T \mathbf{1}}{Y_L^T Y_R}$ and the notation $a_N = \mathcal{O}(b_N)$ means there exists a number $c_0 > 0$ and an integer $N_0 > 0$ such that $|a_N| \leq c_0|b_N|$ $\forall N \geq N_0$. This also implies the convergence to a geometric sequence:

$$\lim_{N\to\infty} \frac{\pi_0^T M^N \mathbf{1}}{\pi_0^T M^{N-1}\mathbf{1}} = \lambda_1 \tag{22}$$

This theorem is widely used, since for finite-state Markov chains, since $P_N = \pi_0^T M^{N-1}\mathbf{1}$, where π_0 is the initial probability vector and M is a submatrix of the probability transition matrix of the Markov chains. For continuous-state Markov Chains, one can approximate P_N by gridding the continuous state. This approach is computationally tractable for low-order systems. Sufficiently many grids can be used to obtain a reasonably accurate result of P_N. For higher order processes, the approximation technique will not be practical because thorough gridding in all dimensions of the random variable will require unreasonable computation time.