2008

Towards a Metric for the Assessment of Safety Critical Control Systems

Oscar R. Gonzalez

Jorge R. Chavez-Fuentes

W. Steven Gray

AIAA Guidance, Navigation and Control Conference and Exhibit
18 - 21 August 2008, Honolulu, Hawaii

AIAA 2008-6804

Towards a Metric for the Assessment of Safety Critical Control Systems

Oscar R. González*, Jorge R. Chávez-Fuentes† and W. Steven Gray‡
Old Dominion University, Norfolk, VA 23529-0246, U.S.A.

There is a need for better integration of the fault tolerant and the control designs for safety critical systems such as aircraft. The dependability of current designs is assessed primarily with measures of the interconnection of fault tolerant components: the reliability function and the mean time to failure. These measures do not directly take into account the interaction of the fault tolerant components with the dynamics of the aircraft. In this paper, a first step to better integrate these designs is made. It is based on the observation that unstable systems are intrinsically unreliable and that a necessary condition for reliability is the existence of a stabilizing control law that depends on the interconnection of the working fault tolerant components. Since operation of a fault tolerant interconnection of digital computers in a harsh environment can result in transient errors, a methodology to analyze the mean square stability of the fault tolerant closed-loop system is presented. A definition for mean square stabilizability is then used to introduce the new dynamical system reliability concept. An example illustrates the effect on mean square stability of several fault tolerant design choices and illustrates possible dynamical system reliability plots.

I. Introduction

EVERY safety critical control system application uses fault tolerant technology to minimize the effect of faults and increase the reliability, dependability, availability, and maintainability of the system.[1–4] The designers of such systems consider several performance metrics to arrive at a final design employing passive and/or active fault tolerant technologies. Passive technologies include replicated sensors, actuators, and processing elements for control system computations and supervisory control; while active fault tolerant technologies include fault detection, isolation and reconfiguration. Assessment of these passive fault tolerant technologies is done starting with fault tolerant metrics of the individual replicated components and deriving a metric for their interconnection. The well-known metrics include the reliability function (probability that a component is operating correctly beyond a given time) and the mean time to failure (MTTF) of the component. There is no well-established assessment of the active fault tolerant technologies nor of the hybrid closed-loop system that includes either type of fault tolerant technology and a continuous-time process. In Refs. 5–8, the closed-loop system with passive fault tolerant components is assumed to have two failure states: failed safe and failed unsafe. Failures are triggered by a fault that led to an error in the operation of the system, which resulted in a significant deviation of the system operation. Additional failure states are needed if active fault tolerant components are used, corresponding to the detected or undetected failure states. The main proposed metric for comparison of these systems is the mean time to fail unsafe (MTTFU). The system safety metric (probability that the system is operational or failed safe at a given time) is also used for comparison in Refs. 6, 7. Fundamentally, these metrics are based on the structure of the interconnection of the fault tolerant components and not on the dynamical system and excitation characteristics. A few metrics based on the response of the dynamical system are available. In Refs. 9, 10, the mean first passage

*The corresponding author. Associate Professor, Systems Research Laboratory, Department of Electrical and Computer Engineering, phone/fax: 757-683-4966/3220, gonzalez@ece.odu.edu
†Ph.D. student, Systems Research Laboratory, Department of Electrical and Computer Engineering, jchav004@odu.edu
‡Associate Professor, Systems Research Laboratory, Department of Electrical and Computer Engineering, gray@ece.odu.edu

time for the response process vector is considered. In their applications, a well-defined boundary for unsafe failure exists.

This paper presents a dynamical system reliability definition for safety critical closed-loop systems that use passive fault tolerant technologies. It is clear that if a closed-loop system is unstable in some sense then it should be considered to be in the fail unsafe state before any of its responses exceed a safe region of operation or while all its components are working as intended; *unstable systems are intrinsically unreliable from the systems point of view*. Of course, safety critical closed-loop systems are designed to be stable within their domain of attraction, which includes the expected failure rates and the associated modes of operation. However, a stable closed-loop system may become unstable if a drastic change takes place such as the failure of one or more components that prevent the system from maintaining stability, resulting in loss of control. Stability can also be lost if the expected failure rates change due to the introduction of the safety critical system into a harsh environment. This is analogous to stability of nonlinear systems, where the change of an input can change the equilibria and stability properties. In this case, the random failures can be considered to be an input. Since the failures occur randomly, mean square stability (MSS) of the closed-loop system is analyzed. So, closed-loop systems that are not MSS or abruptly stop being MSS are in the fail unsafe state. Systems with responses outside a safe region of operation are also in the fail unsafe state. This region can be characterized with a cost function as done in Refs. 11, 12 or by specifying performance boundaries as in Refs. 9, 10, 13. Thus, *a necessary condition for system reliability is MSS*. From a practical point of view, if an abrupt change has taken place, a more general necessary condition for dynamical system reliability is the existence of a reconfigurable control law to return operation to at least a fail safe state. In this paper, a type of mean square stabilizability will be presented that makes it possible to determine if the closed-loop has recovered. A relation between system reliability and stabilizability was first articulated in Ref. 14.

To analyze MSS of passive fault tolerant interconnected closed-loop systems, a methodology based on Markov jump linear systems[15] will be presented. If certain technical assumptions are satisfied then it is possible to analyze MSS of an equivalent jump linear system driven by a homogeneous Markov process. In this case, it is known that MSS implies stability in the mean, almost sure stability, and exponential mean square stability.[16,17] For simplicity, in this paper only fault tolerance of the controller is considered. Failures of each fault tolerant component will be characterized with independent homogeneous Markov chains with only two states: not failed and failed. Failure of the interconnection of fault tolerant components is characterized with a structure function.[18] These models result in a closed-loop system realization that is randomly switched. Since, in general, the switching process is not a homogenous Markov chain,[19] a joint process is introduced that results in an equivalent Markov jump linear system. For comparison of different fault tolerant implementations, the MSS stability boundary is computed as a function of the transition probability matrix parameters. The direct effect of faults in the communication system is not considered. An investigation that takes this into account and analyzes its effect on a control system performance measure is Ref. 20.

The rest of the paper is organized as follows. In Section II the fault tolerant architecture to be considered is presented. The statistical nature of the stochastic processes associated with the interconnection of fault tolerant components is presented in Section III. This section also presents an equivalent Markov jump linear system that is suitable for stability analysis. The mean square stability definition and test as well as the dynamical system reliability definition are introduced in Section IV. In Section V, the methodology is applied to a simplified model of an AFTI-F16 aircraft for illustration.

II. Fault Tolerant Architecture

This section describes an N-Modular Redundant (NMR) implementation of a control law for a unity feedback closed-loop system (see Figure 1). Let the underlying probability space be $(\Omega, \mathcal{F}, \Pr)$. Since stability of the randomly switched sampled-data system is equivalent to the stability of a discretized realization,[21] let the controlled process have the following discretized state space realization

$$\begin{aligned} \boldsymbol{x}_{\mathrm{p}}(k+1) &= A_{\mathrm{p}}\boldsymbol{x}_{\mathrm{p}}(k) + B_{\mathrm{p}}\boldsymbol{u}(k) \\ \boldsymbol{y}_{\mathrm{p}}(k) &= C_{\mathrm{p}}\boldsymbol{x}_{\mathrm{p}}(k), \end{aligned} \tag{1}$$

where $\boldsymbol{x}_{\mathrm{p}}(k) \in \mathbb{R}^{n_{\mathrm{p}}}$ is the plant's state vector, $\boldsymbol{y}_{\mathrm{p}}(k) \in \mathbb{R}^{m}$ is the plant's output, A_{p}, B_{p}, and C_{p} are matrices with appropriate dimensions. Boldfaced characters denote a random variable or process. The control law is implemented in N Processing Elements (PE's) followed by an NMR logic circuit that calculates the actual

control command, $\boldsymbol{u}(k)$, sent to the actuators. The PE's satisfy the following simplifying assumptions.

Assumption 1. *Each PE is identical and runs the same nominal control law given by*

$$\begin{aligned} \boldsymbol{x}_{\mathrm{c}}(k+1) &= A_{\mathrm{c}}\boldsymbol{x}_{\mathrm{c}}(k) + B_{\mathrm{c}}\boldsymbol{e}(k) \\ \boldsymbol{y}(k) &= C_{\mathrm{c}}\boldsymbol{x}_{\mathrm{c}}(k), \end{aligned} \tag{2}$$

where $\boldsymbol{x}_{\mathrm{c}}(k) \in \mathbb{R}^{n_{\mathrm{c}}}$ is the controller's state vector, $\boldsymbol{w}(k) \in \mathbb{R}^m$ represents a random additive input perturbation, $\boldsymbol{e}(k) = r(k) + \boldsymbol{w}(k) - \boldsymbol{y}_{\mathrm{p}}(k)$ is the controller's input, $r(k) \in \mathbb{R}^m$ is a deterministic reference input, $\boldsymbol{y}(k) \in \mathbb{R}^m$ is the controller's output, and A_{c}, B_{c}, and C_{c} are matrices with appropriate dimensions.

Assumption 2. *Each PE belongs to a different fault containment region.*

Assumption 3. *Each PE is implemented in a fail silent control computer that includes self-checking such that when an internal upset is detected, the PE stops transmitting data and broadcasts a failed message. The NMR logic circuit is assumed not to fail. The fail silent detection is also assumed not to fail, that is, it has 100% coverage.*

Assumption 4. *The transition between failed and not failed states of each PE is characterized by a homogenous Markov chain $\{\boldsymbol{z}_i(k), k \in \mathbb{Z}^+\}$, $i = 1, \ldots, N$ taking values in $\{0, 1\}$, where $\mathbb{Z}^+$ denotes the non-negative integers. The not failed state is denoted by 0 and the failed one by 1. The stochastic processes $\boldsymbol{z}_1(k), \ldots, \boldsymbol{z}_N(k)$ are assumed to be independent.*

As seen in Figure 1, each PE controller has the same input $\boldsymbol{e}(k)$. Let the ideal output of the control law (2) at time k be $\boldsymbol{y}(k)$. Then the output of the i-th PE controller satisfies $\boldsymbol{y}_i(k) = \boldsymbol{y}(k)$ when $\boldsymbol{z}_i(k) = 0$; otherwise, by Assumption 3, there is no data present in its output. At each time k, the random variables $\boldsymbol{z}_i(k)$ are indicators of the event $\{\boldsymbol{y}_i(k) \neq \boldsymbol{y}(k)\}$, that is, the event that $\{\boldsymbol{y}_i(k)$ is not correct$\}$. Alternatively, the random variable $\boldsymbol{z}_i(k)$ is an indicator of the availability of the i-th PE controller. In Figure 1, the indicator random variables associated with the output of a device are denoted with dashed lines. The NMR logic block characterizes the operation used to compute the control command $\boldsymbol{u}(k)$. In this paper, the NMR logic block satisfies the following assumption.

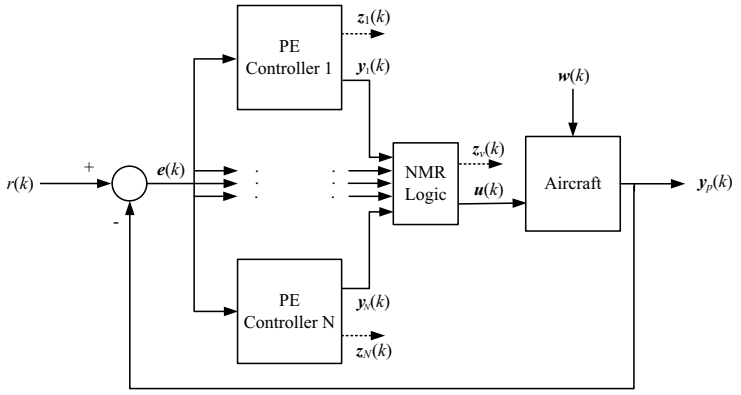

Figure 1. Block diagram of a safety critical closed-loop system with an N-modular redundant controller architecture.

Assumption 5. *At each time k, the NMR logic circuit implements the following input/output relation*

$$\boldsymbol{u}(k) = g_{\mathrm{L}}\big(\boldsymbol{y}_1(k), \ldots, \boldsymbol{y}_N(k)\big),$$

where $g_{\mathrm{L}} : \{\mathbb{R}^m\}^N \to \mathbb{R}^m$ with $\{\mathbb{R}^m\}^N \triangleq \underbrace{\mathbb{R}^m \times \cdots \times \mathbb{R}^m}_{N \text{ times}}$. By Assumption 3, the mapping g_{L} results in the following simple assignment

$$\boldsymbol{u}(k) = \begin{cases} \boldsymbol{y}(k) & : \text{ if there exists at least one } i \text{ such that } \boldsymbol{y}_i(k) = \boldsymbol{y}(k), \\ - & : \text{ otherwise} \end{cases}$$

where the dash indicates that the NMR logic circuit was unable to determine a value for the control command. The control command is undefined in this case.

Assumption 5 simply says that because of the assumed 100% coverage in Assumption 3, only one PE needs to be in operation as in a parallel interconnection of the PE's. If the coverage assumption is relaxed, the mapping g_{L} could implement one of several possible voting algorithms such as majority or weighted average.[22–24] So, in general, the correctness of $\boldsymbol{u}(k)$ is characterized by a transformation of the random processes indicating if each PE is working properly. This transformation, found from reliability block diagrams or fault tree analysis, is called the structure function and is defined next.

Definition 1. *A structure function, ϕ, for the fault tolerant interconnection of PE's in Figure 1 is a memoryless, binary-valued mapping $\phi : \{0,1\}^N \to \{0,1\}$ given by*

$$\boldsymbol{z}_v(k) = \phi(\boldsymbol{z}(k)) = \begin{cases} 0 & : \ \boldsymbol{u}(k) = \boldsymbol{y}(k), \\ 1 & : \ \boldsymbol{u}(k) \text{ is undefined}, \end{cases} \tag{3}$$

where $\boldsymbol{z}(k) = \big(\boldsymbol{z}_1(k), \ldots, \boldsymbol{z}_N(k)\big)$.

In (3), $\boldsymbol{z}_v(k)$ is the indicator random variable associated with $\boldsymbol{u}(k)$, the output of the NMR logic circuit. The structure function for many interconnections is known.[8] A general class of such interconnections are referred to as α-out-of-N, where α PE's need to function for correct operation of the interconnected system. The specific case considered in this paper is a 1-out-of-N interconnected system.

The possible ambiguity caused by $\boldsymbol{u}(k)$ being undefined is resolved in practice in the actuators. Thus, two types of actuators are considered: memoryless actuators and actuators with memory. The actuators satisfy the following assumptions.

Assumption 6. *Memoryless actuators assume a zero command when no data is received. The effective control input seen by them is a function of $\boldsymbol{z}_v(k)$:*

$$\begin{aligned} \boldsymbol{u}(k) &= \boldsymbol{u}_{\boldsymbol{z}_v(k)}(k) \\ &= (1 - \boldsymbol{z}_v(k))\, \boldsymbol{y}(k). \end{aligned}$$

Assumption 7. *Actuators with memory belong to a class of smart actuators. When no data is received, these actuators use the previous control command. The effective control input is*

$$\begin{aligned} \boldsymbol{u}(k) &= \boldsymbol{u}_{\boldsymbol{z}_v(k)}(k) \\ &= (1 - \boldsymbol{z}_v(k))\, \boldsymbol{y}(k) + \boldsymbol{z}_v(k) \boldsymbol{y}(k-1). \end{aligned}$$

Now it is possible to develop the dynamical closed-loop models that are needed to study MSS of this class of fault tolerant closed-loop systems. A realization of the closed-loop system in Figure 1 is

$$\begin{aligned} \boldsymbol{x}_{\mathrm{CL}}(k+1) &= A_{\boldsymbol{z}_v(k)} \boldsymbol{x}_{\mathrm{CL}}(k) + B_{\boldsymbol{z}_v(k)} (r(k) + \boldsymbol{w}(k)) \\ \boldsymbol{y}_{\mathrm{CL}}(k) &= C_{\boldsymbol{z}_v(k)} \boldsymbol{x}_{\mathrm{CL}}(k), \end{aligned} \tag{4}$$

where $\boldsymbol{y}_{\mathrm{CL}}(k) = \boldsymbol{y}_{\mathrm{p}}(k)$ and, for example, for the simplex architecture ($N = 1$) with memoryless actuators,

$$A_{\boldsymbol{z}_v(k)} = \begin{cases} A_0 = \begin{bmatrix} A_{\mathrm{p}} & -B_{\mathrm{p}} C_{\mathrm{c}} \\ B_{\mathrm{c}} C_{\mathrm{p}} & A_{\mathrm{c}} \end{bmatrix} & : \ \boldsymbol{z}_v(k) = 0, \\ A_1 = \begin{bmatrix} A_{\mathrm{p}} & 0 \\ 0 & 0 \end{bmatrix} & : \ \boldsymbol{z}_v(k) = 1, \end{cases}$$

and $\boldsymbol{x}_{\mathrm{CL}}(k) = [\boldsymbol{x}_{\mathrm{p}}^{\mathrm{T}}(k) \ \boldsymbol{x}_{\mathrm{c}}^{\mathrm{T}}(k)]^{\mathrm{T}} \in \mathbb{R}^n$ with $n = n_{\mathrm{p}} + n_{\mathrm{c}}$. Only $A_{\boldsymbol{z}_v(k)}$ is given, since it is the only matrix needed for stability analysis. When the actuators have memory, the closed-loop system can be augmented with an additional state vector that remembers the previous value of the controller's state vector. So, the

state vector in (4) is $\boldsymbol{x}_{\mathrm{CL}}(k) = [\boldsymbol{x}_{\mathrm{p}}^{\mathrm{T}}(k)\ \boldsymbol{x}_{\mathrm{c}}^{\mathrm{T}}(k)\ \boldsymbol{x}_{a}^{\mathrm{T}}(k)]^{\mathrm{T}} \in \mathbb{R}^{n_{\mathrm{p}}+2n_{\mathrm{c}}}$, $\boldsymbol{x}_a(k) = \boldsymbol{x}_{\mathrm{c}}(k-1)$, and

$$A_{\boldsymbol{z}_v(k)} = \begin{cases} A_0 = \begin{bmatrix} A_{\mathrm{p}} & -B_{\mathrm{p}}C_{\mathrm{c}} & 0 \\ B_{\mathrm{c}}C_{\mathrm{p}} & A_{\mathrm{c}} & 0 \\ 0 & I & 0 \end{bmatrix} & : \ \boldsymbol{z}_v(k) = 0, \\ A_1 = \begin{bmatrix} A_{\mathrm{p}} & 0 & -B_{\mathrm{p}}C_{\mathrm{c}} \\ 0 & 0 & I \\ 0 & 0 & I \end{bmatrix} & : \ \boldsymbol{z}_v(k) = 1. \end{cases}$$

The $A_{\boldsymbol{z}_v(k)}$ matrices can be derived in a similar way for other fault tolerant interconnections. The statistical nature of $\boldsymbol{z}(k)$ and $\boldsymbol{z}_v(k)$ will be characterized in the next section. A new joint process is introduced that results in a Markov jump linear system realization equivalent to (4).

III. Markov Jump Linear Systems

The main goal of this section is to develop a Markov Jump Linear System (MJLS) realization of the randomly switched system in (4). First, the statistical nature of $\boldsymbol{z}(k)$ is characterized in Lemma 1. Note that the random processes $\boldsymbol{z}_1(k), \ldots, \boldsymbol{z}_N(k)$ are independent if the random variables at time k are mutually independent for every $k \in \mathbb{Z}^+$.

Lemma 1. *Let $\boldsymbol{z}_1(k), \ldots, \boldsymbol{z}_N(k)$ be independent homogeneous Markov chains on a finite or countable state space $\mathcal{I}_i$ and with transition probability matrices Π_{z_i}, $i = 1, \ldots, N$ over the same probability space $(\Omega, \mathcal{F}, \Pr)$. Then the joint process $\boldsymbol{z}(k) \triangleq (\boldsymbol{z}_1(k), \ldots, \boldsymbol{z}_N(k))$ is a homogeneous Markov chain with state space $\mathcal{I}_1 \times \cdots \times \mathcal{I}_N$ and transition probability matrix*

$$\Pi_z \triangleq \Pi_{z_1} \otimes \cdots \otimes \Pi_{z_N},$$

where $\otimes$ denotes the Kronecker product. The joint process, $\boldsymbol{z}(k)$, is irreducible and aperiodic if each of the Markov chains, $\boldsymbol{z}_i(k)$, satisfies these properties. Moreover, $\boldsymbol{z}(k)$ is recurrent nonnull if invariant distributions exist for $\boldsymbol{z}_i(k)$, $i = 1, \ldots, N$.

Proof. This is a direct generalization of Lemma 7.19 in Ref. 25. □

Even though the random process $\boldsymbol{z}_v(k)$ in (3) is a memoryless, binary-valued function of $\boldsymbol{z}(k)$, a homogeneous Markov process, $\boldsymbol{z}_v(k)$ is not in general a homogeneous Markov process.[19,26,27] To facilitate the MSS analysis of the closed-loop system, an augmented joint process is introduced and shown to be Markov and homogeneous.

Theorem 1. *Let $\boldsymbol{z}_1(k), \ldots, \boldsymbol{z}_N(k)$ be independent homogeneous Markov chains on $\{0, 1\}$ with transition probability matrices Π_{z_i}, $i = 1, \ldots, N$. For each $k \in \mathbb{Z}^+$ let $\boldsymbol{z}(k) \triangleq (\boldsymbol{z}_1(k), \ldots, z_N(k))$ and let $\boldsymbol{z}_v(k)$ be given by (3). Then the joint process $\boldsymbol{\rho}(k) \triangleq (\boldsymbol{z}(k), \boldsymbol{z}_v(k))$ is a homogeneous Markov chain, and its state space can be reduced to a proper subset of $\{0, 1\}^{N+1}$ such that its transition probability matrix satisfies $\Pi_\rho = \Pi_z$. The joint process, $\boldsymbol{\rho}(k)$ is ergodic if invariant distributions exist for $\boldsymbol{z}_i(k)$ $(i = 1, \ldots, N)$, and they are irreducible and aperiodic.*

Proof. By Lemma 1, $\boldsymbol{z}(k)$ is a Markov chain with transition probability matrix given by Π_z. By Theorems 3.2–3.4 in Ref. 28, the following σ-algebra relationship holds $\sigma(\boldsymbol{\rho}(k), \ldots, \boldsymbol{\rho}(0)) = \sigma(\boldsymbol{z}(k), \ldots, \boldsymbol{z}(0))$, since $\boldsymbol{\rho}(k) = (\boldsymbol{z}(k), \phi(\boldsymbol{z}(k)))$, and ϕ is a binary-valued function of $\boldsymbol{z}(k)$. To simplify the notation, for any $k \in \mathbb{Z}^+$, denote the events $\{\boldsymbol{\rho}(k) = \rho(k)\}$, $\{\boldsymbol{\rho}(k-1) = \rho(k-1), \ldots, \boldsymbol{\rho}(0) = \rho(0)\}$, and $\{\boldsymbol{z}(k-1) = z(k-1), \ldots, \boldsymbol{z}(0) = z(0)\}$ by $\{\rho(k)\}$, $\{\rho(k-1), \ldots, \rho(0)\}$, and $\{z(k-1), \ldots, z(0)\}$, respectively. Since ϕ is a memoryless mapping, there is a one-to-one mapping between $\boldsymbol{\rho}(k)$ and $\boldsymbol{z}(k)$. Thus the number of states that $\boldsymbol{\rho}(k)$ can take with non-zero probability is the same as the number of states $\boldsymbol{z}(k)$ can assume, that is,

2^N. Thus, $\Pr\{(z(k), \phi(z(k)))\} = \Pr\{z(k)\}$. Now, since $z(k)$ is Markov

$$\begin{aligned}
&\Pr\{\rho(k)|\{\rho(k-1), \ldots, \rho(0)\}\} \\
&\quad = \Pr\{(z(k), \phi(z(k)))|\{z(k-1), \ldots, z(0)\}\} \\
&\quad = \Pr\{z(k)|\{z(k-1)\}\} \\
&\quad = \Pr\{\rho(k)|\{\rho(k-1)\}\}.
\end{aligned}$$

Therefore, $\boldsymbol{\rho}(k)$ is Markov, and it has the same transition probability matrix as $\boldsymbol{z}(k)$. Finally, since $\boldsymbol{\rho}(k)$ is completely characterized by $\boldsymbol{z}(k)$, Lemma 1 also determines whether it is irreducible, aperiodic, and recurrent nonnull and, hence, whether it is ergodic or not. Note that recurrent states in a Markov chain with finite states can only be nonnull. □

The Markov chain $\boldsymbol{\rho}(k)$ can be used to define the following MJLS

$$\begin{aligned}
\boldsymbol{x}_{\mathrm{CL}}(k+1) &= A_{\boldsymbol{\rho}(k)}\boldsymbol{x}_{\mathrm{CL}}(k) + B_{\boldsymbol{\rho}(k)}(r(k) + \boldsymbol{w}(k)) \\
\boldsymbol{y}_{\mathrm{CL}}(k) &= C_{\boldsymbol{\rho}(k)}\boldsymbol{x}_{\mathrm{CL}}(k),
\end{aligned} \tag{5}$$

which is selected to be *model equivalent* to the randomly switched system in (4), that is, for each $k \in \mathbb{Z}^+$ $A_{\boldsymbol{\rho}(k)} \equiv A_{\boldsymbol{z}_v(k)}$, $B_{\boldsymbol{\rho}(k)} \equiv B_{\boldsymbol{z}_v(k)}$, and $C_{\boldsymbol{\rho}(k)} \equiv C_{\boldsymbol{z}_v(k)}$.[29] Therefore, if (4) and (5) have the same initial conditions and input processes, their state and output processes will be the same.

The next section reviews MSS for MJLS and presents definitions for mean square stabilizability and dynamical system reliability.

IV. MSS and Dynamical System Reliability

This section starts with a summary of a well-known necessary and sufficient mean square stability condition for a class of MJLS. An in-depth analysis can be found in Ref. 15.

Consider the following MJLS

$$\boldsymbol{x}(k+1) = A_{\boldsymbol{\rho}(k)}\boldsymbol{x}(k) + B_{\boldsymbol{\rho}(k)}\boldsymbol{w}(k), \quad \boldsymbol{x}(0) = \boldsymbol{x}_0, \tag{6}$$

where $\boldsymbol{x}(k) \in \mathbb{R}^n$; $\boldsymbol{\rho}(k)$ is a homogeneous, discrete-time Markov chain that takes values in the finite set $\mathcal{I}_{l_\rho} = \{0, \ldots, l_\rho - 1\}$; $A_i \in \mathbb{R}^{n\times n}$ for all $i \in \mathcal{I}_{l_\rho}$; and $\boldsymbol{x}(0)$ is a random vector with finite second moment that is independent of $\boldsymbol{\rho}(k)$ for $k \geq 0$. Also, let $\Pi_\rho = [\pi_{ij}]$, $\pi(k)$, and $\pi(0) = \pi_0$ represent the transition probability matrix, the distribution at time k, and the initial distribution of $\boldsymbol{\rho}(k)$, respectively. When $\boldsymbol{w}(k) \neq 0$, the following extra assumption is needed.

Assumption 8. *When (6) is not autonomous, i.e., $\boldsymbol{w}(k) \neq 0$ for some $k \geq 0$, assume that $\boldsymbol{\rho}(k)$ is ergodic (i.e., it has a single ergodic class), that the processes $\boldsymbol{\rho}(k)$ and $\boldsymbol{w}(k)$ are independent of each other, and $\boldsymbol{w}(k)$ is independent of the initial state $\boldsymbol{x}(0)$ for $k \geq 0$. Furthermore, $\boldsymbol{w}(k)$ is considered to be a second-order, independent, wide sense stationary sequence of random variables.*

The following stability definition is standard in the literature.

Definition 2. *The equilibrium point at 0 of system (6) (or simply, system (6)) is called mean square stable if there exists a nonnegative constant α such that for every value of the initial condition $\boldsymbol{x}_0$ and every initial distribution π_0 of $\boldsymbol{\rho}(k)$ it follows that $\lim_{k\to\infty} E\{\|\boldsymbol{x}(k)\|^2\} = \alpha$. If $\boldsymbol{w}(k) = 0$ for $k \geq 0$ then $\alpha = 0$.*

The main mean square stability test for an MJLS follows.[15]

Theorem 2. *System (6) is mean square stable if and only if the spectral radius of $\mathcal{A}$ is less than one, where*

$$\mathcal{A} \triangleq (\Pi_\rho^T \otimes I_{n^2})\,\mathrm{diag}(A_0 \otimes A_0, \ldots, A_{l_\rho-1} \otimes A_{l_\rho-1}).$$

Note that whenever (6) is not autonomous, Assumption 8 is needed to ensure the uniqueness of the limit

$$\lim_{k\to\infty} E\{\|\boldsymbol{x}(k)\|^2\}$$

in Definition 2. Ergodicity of $\boldsymbol{z}(k)$, however, is not required when $\boldsymbol{w}(k) = 0$ for all $k \geq 0$.[15,16]

From the control point of view it is important to determine if there exists a reconfigurable control law after a failure or failures occur. Mean square stabilizability for the class of systems presented here is defined next.

Definition 3. *The aircraft dynamics in (1) are said to be mean square stabilizable if the interconnection of N PE's running the control law in (2) stabilizes (1) in the mean square sense with the available control input* $\boldsymbol{u}_{\boldsymbol{\rho}(k)}(k)$.

Mean square stabilizability will be denoted by $\mathrm{MSS}_{\boldsymbol{\rho}(k)}$ to emphasize the dependence on the working interconnected fault tolerant components. Next a dynamical system reliability definition is introduced. First, a standard definition for the reliability function of a fault tolerant component is given. It is followed by the definition of a reliability function for an interconnected system of fault tolerant components.

Definition 4. *Let the random variable* $\boldsymbol{q}_i$ *be the time to failure of the i-th fault tolerant component that starts operation at time* $t = 0$, $i = 1, \ldots, N$. *The reliability function of the i-th component is*

$$R_i(t) \triangleq \Pr\{\boldsymbol{q}_i > t\} = 1 - F_i(t),$$

where $F_i(t) = \Pr\{\boldsymbol{q}_i \leq t\}$ *is the distribution of the continuous-time random variable* $\boldsymbol{q}_i$ *and* $t \in \mathbb{R}^+$ *the nonnegative reals.*

The closed-loop system in Figure 1 is a sampled-data system with continuous-time aircraft dynamics and a digital controller implementation. Let T denote the sample period. Note that for every $t \in \mathbb{R}^+$ there exists a $k \in \mathbb{Z}^+$, corresponding to the previous sample instant, such that

$$kT \leq t < (1 + \Delta)kT,$$

where $0 < \Delta < 1$. For fault tolerant sampled-data systems, the conditional reliability may be more relevant if it is known that the component was working at kT. This conditional reliability satisfies

$$R_i(\Delta kT, kT) \triangleq \frac{R_i((1 + \Delta)kT)}{R_i(kT)}.$$

If the component lifetimes are exponentially distributed then it is known that $R_i(\Delta kT, kT) = R_i(\Delta kT)$.[8]

Definition 5. *The reliability function of an interconnected system of N fault tolerant components is*

$$R(t) \triangleq \Pr\{\boldsymbol{q} > t\} = 1 - F(t),$$

where $\boldsymbol{q}$ *is the lifetime of the fault tolerant system, and* $F(t)$ *is its distribution function.*

The lifetime of the interconnected system, $\boldsymbol{q}$, can be found from the lifetime of the components, that is, there is a mapping that depends on the structure of the interconnection from $(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_N)$ to $\boldsymbol{q}$.

For a dynamical closed-loop system the following definition of reliability integrates the dynamical system and the fault tolerant system.

Definition 6. *The dynamical system reliability is*

$$R_S(t) = \Pr\{\boldsymbol{q} > t\}\mathbf{1}_{\{MSS_{\boldsymbol{\rho}(k)}\}},$$

where $\boldsymbol{q}$ *is the time to failure of the interconnection of fault tolerant components, and* $\mathbf{1}_{\{\cdot\}}$ *is the indicator of the event* $\{MSS_{\boldsymbol{\rho}(k)}\}$ *that the aircraft dynamics are mean square stabilizable with the nominal controller and the operational fault tolerant interconnected system.*

This is the probability of the time to failure of the interconnection of the fault tolerant components times the indicator of the event that there exists a mean square stabilizing control law with working components.

V. Example

This section presents a simulation study to illustrate the effect of several fault tolerant architecture choices on the MSS and the dynamical system reliability. The example uses the simplified longitudinal dynamics of the AFTI-F16 aircraft given in Ref. 30. The aircraft model has four states (change in speed, angle of attack, pitch rate, and pitch angle). In this paper, it is controlled by an observer-based digital regulator given by

$$\begin{aligned} \boldsymbol{x}_{\mathrm{c}}(k+1) &= A_{\mathrm{p}}\boldsymbol{x}_{\mathrm{c}}(k) + B_{\mathrm{p}}\boldsymbol{u}(k) + F(\boldsymbol{y}_{\mathrm{p}}(k) - C_{\mathrm{p}}\boldsymbol{x}_{\mathrm{c}}(k)) \\ \boldsymbol{y}_{\mathrm{c}}(k) &= \boldsymbol{x}_{\mathrm{c}}(k), \end{aligned}$$

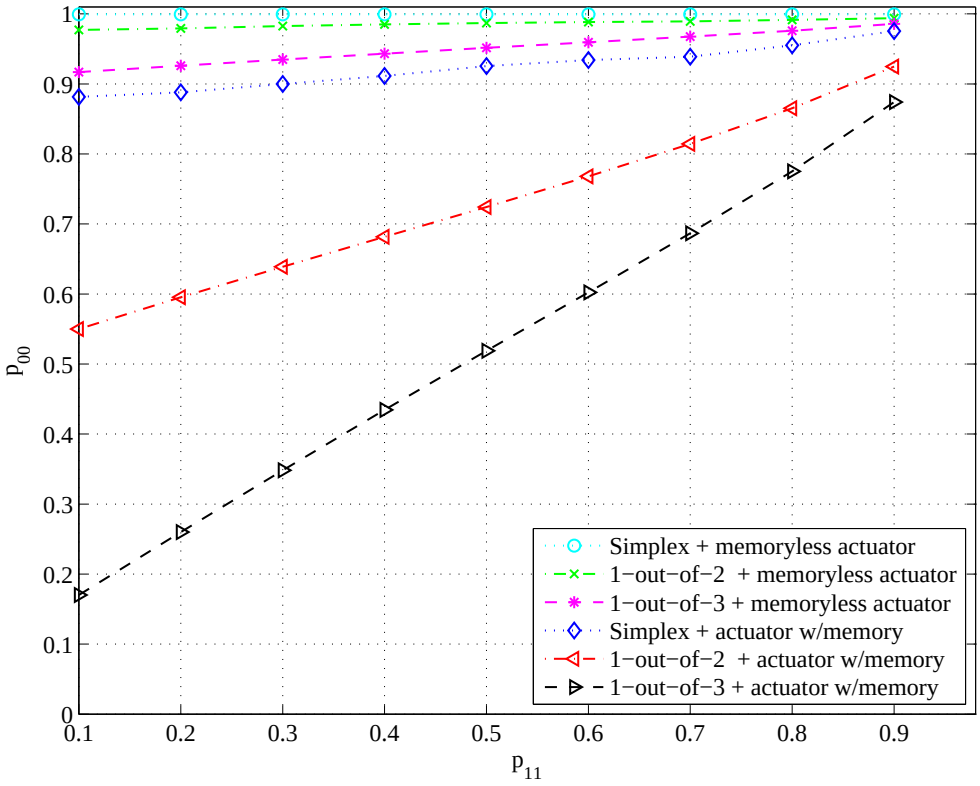

Figure 2. Plots of the stability boundaries for the AFTI-F16 example. In each case the stable region is above the boundary.

where $\boldsymbol{u}(k) = r(k) - K\boldsymbol{x}_{\rm c}(k)$, and F is the observer gain matrix. Here $\boldsymbol{x}_{\rm c}(k) = \hat{\boldsymbol{x}}_{\rm p}(k)$ is an estimate of the aircraft's state vector. The sampling period was set to $T = 0.004$ sec., the nominal continuous-time closed-loop poles were placed at $\{-0.2 \pm j0.9798, -0.01 \pm j0.0995\}$, and the observer's discrete-time poles were chosen to be five times faster than the plant's closed-loop poles.

The observer-based digital regulator was implemented in 1 or 2 or 3 PE's. The failures of each PE were characterized with independent homogenous Markov chains with two states: 0 (operational) and 1 (failure) with transition probability matrix

$$\Pi = \begin{bmatrix} p_{00} & 1-p_{00} \\ 1-p_{11} & p_{11} \end{bmatrix},$$

where p_{00} is the probability that the PE stays in the operational mode and p_{11} the probability of staying in the failure mode. To complete the description of the fault tolerant interconnected system, a 1-out-of-N NMR logic circuit, satisfying the assumptions in Section II was implemented. For each number of PE's, a memoryless actuator and actuator with memory case is considered, resulting in six cases. For these cases, no NMR logic failures are considered.

Now that the closed-loop system and the fault tolerant interconnected systems have been defined, Theorem 2 is used to sweep the transition probability parameters (p_{00} and p_{11}) to find the stability boundaries for each architecture choice. The results are shown in Figure 2, which gives the tradeoff between the persistence of each PE staying in the operational state (p_{00}) vs. the persistence of each PE staying in a failure mode (p_{11}). According to Figure 2, the best of the considered architectures is the 1-out-of-3 one with actuators with memory, since it has the largest stability region. The worst architecture is the simplex one with memoryless actuators. The stability boundaries corresponding to the other architectures fall in between as intuitively expected.

To illustrate the use of the dynamical system reliability, note that for a 1-out-of-N interconnection of fault tolerant components, the reliability function of the interconnected system is

$$R(t) = 1 - \prod_{i=1}^{N}(1 - R_i(t)),$$

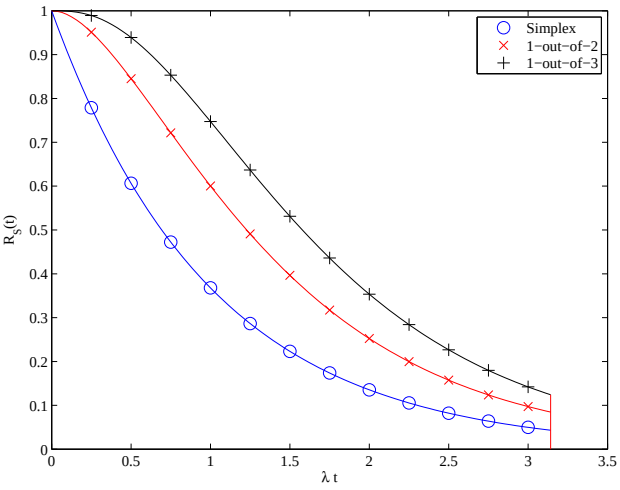

Figure 3. The dynamical system reliability plots for 1-out-of-N interconnections for $N = 1, 2, 3$. The dynamical system ceases to be stabilizable when $\lambda t = \pi$.

where $R_i(t)$, $i = 1, \ldots, N$ is the reliability function of the i-th component.[8] If the component reliabilities are the same and given by $R_i(t) = e^{-\lambda t}$, and if the closed-loop cannot be stabilized after an abrupt change at $\lambda t = \pi$, then an illustration of the dynamical system reliability is given in Figure 3. For other α-out-of-N interconnections or when the coverage of the failures is not 100%, it is expected that the dynamical system reliability functions will have different abrupt jumps to zero.

VI. Conclusions

In this paper, a methodology to analyze the MSS of a class of fault tolerant computer control systems was presented. The main contribution is the characterization of the joint process used to drive the model equivalent Markov jump linear system. It was shown that it is a homogeneous Markov process with the same transition probability matrix as the joint process of the N indicator random variables characterizing the availability of the PE's. Current research directions include determining the conditions under which the 2-state indicator random variable $\boldsymbol{z}_v(k)$ can be used to drive another model equivalent Markov jump linear system. By using $\boldsymbol{z}_v(k)$ instead of $\boldsymbol{\rho}(k)$, tests with much lower dimensional matrices can be used for MSS and control system performance. Another contribution of this paper is the introduction of a mean square stabilizability condition used to define a new dynamic system reliability function. Future work will show how to use this function in optimization of control systems for improved reliability.

Acknowledgments

This research was supported by the NASA Langley Research Center under grant NNX07AD52A.

References

[1]Avižienis, A., Laprie, J.-C., Randell, B., and Landwehr, C., "Basic Concepts and Taxonomy of Dependable and Secure Computing," *IEEE Trans. Dependable Secure Comput.*, Vol. 1, No. 1, 2004, pp. 11–33.

[2]Frank, P. M., "Fault Diagnosis in Dynamic Systems Using Analytical and Knowledge-based Redundancy – A Survey and Some New Results," *Automatica*, Vol. 26, No. 3, 1990, pp. 459–474.

[3]Johnson, B. W., *Design and Analysis of Fault Tolerant Digital Systems*, Addison-Wesley Publishing Co., Reading, MA, 1989.

[4]Mahmoud, M., Jiang, J., and Zhang, Y., *Active Fault Tolerant Control Systems – Stochastic Analysis and Synthesis*, Vol. 287 of *Lecture Notes in Control and Information Sciences*, Springer-Verlag, Berlin, 2003.

[5]Bukowski, J. V. and Goble, W. M., "Defining Mean-time-to-failure in a Particular Failure-state for Multi-failure-state Systems," *IEEE Trans. Rel.*, Vol. 50, No. 2, 2001, pp. 221–228.

[6]DeLong, T. A., Smith, D. T., and Johnson, B. W., "Dependability Metrics to Assess Safety-critical Systems," *IEEE Trans. Rel.*, Vol. 54, No. 3, 2005, pp. 498–505.

[7]Levitin, G., Zhang, T., and Xie, M., "State Probability of a Series-parallel Repairable System with Two-types of Failure States," *Int. J. of Systems Science*, Vol. 37, No. 14, 2006, pp. 1011–1020.

[8]Trivedi, K. S., *Probability and Statistics with Reliability, Queueing and Computer Science Applications*, John Wiley & Son, Inc., New York, NY, 2nd ed., 2002.

[9]Richard V. Field, J. and Bergman, L. A., "Reliability-based Covariance Control Design," *Proc. 1997 American Control Conf.*, Albuquerque, NM, 1997, pp. 11–15.

[10]Song, J. and Kiureghian, A. D., "Joint First-passage Probability and Reliability of Systems under Stochastic Excitation," *ASCE J. Engr. Mech.*, Vol. 132, No. 1, 2006, pp. 65–77.

[11]Li, H., Zhao, Q., and Yang, Z., "Reliability Monitoring of Fault Tolerant Control Systems with Demonstration on an Aircraft Model," *J. Contr. Sci. Eng.*, Vol. 2008, Article ID 265189, 2008.

[12]Wu, N. E., "Coverage in Fault-tolerant Control," *Automatica*, Vol. 40, 2004, pp. 537–548.

[13]Tejada, A., González, O. R., and Gray, W. S., "Asymptotic and Mean Square Stability Conditions for Hybrid Jump Linear System with Performance Supervision," *Proc. 2005 American Control Conf.*, Portland, OR, 2005, pp. 569–574.

[14]Birdwell, J. D., Castanon, D. A., and Athans, M., "On Reliable Control System Designs," *IEEE Trans. Syst., Man, Cybern.*, Vol. 16, No. 5, 1986, pp. 703–711.

[15]Costa, O. L. V., Fragoso, M. D., and Marques, R. P., *Discrete-time Markov Jump Linear Systems*, Springer, London, 2005.

[16]Fang, Y., Loparo, K. A., and Feng, X., "Stability of Discrete-time Jump Linear Systems," *J. of Mathematical Systems, Estimation and Control*, Vol. 5, 1995, pp. 275–321.

[17]Ji, Y. and Chizeck, H. J., "Jump Linear Quadratic Gaussian Control: Steady State Solution and Testable Conditions," *Control Theory Adv. Tech.*, Vol. 6, No. 3, 1990, pp. 289–319.

[18]Barlow, R. E. and Proschan, F., *Statistical Theory of Reliability and Life Testing: Probability Models*, Holt, Rinehart and Winston, New York, NY, 1975.

[19]Gurvits, L. and Ledoux, J., "Markov Property for a Function of a Markov Chain: A Linear Algebra Approach," *Lin. Algebra Appl.*, Vol. 404, July 2005, pp. 85–117.

[20]Gray, W. S., Wang, R., and González, O. R., "A Performance Model for a Distributed Flight Control System Subject to Random Upsets," *Proc. 2008 Conf. on Control Applications*, San Antonio, TX, 2008, pp. 918–923.

[21]González, O. R., Herencia-Zapana, H., and Gray, W. S., "Stochastic Stability of Nonlinear Sampled-data Systems with a Jump Linear Controller," *Proc. 43rd IEEE Conf. on Decision and Control*, Nassau, Bahamas, 2004, pp. 4128–4133.

[22]Latif-Shabgahi, G. R., "A Novel Algorithm for Weighted Average Voting Used in Fault Tolerant Computing Systems," *Microprocessors and Microsystems*, Vol. 28, 2004, pp. 357–361.

[23]Latif-Shabgahi, G. R., Bass, J. M., and Bennet, S., "A Taxonomy for Software Voting Algorithms Used in Safety-critical Systems," *IEEE Trans. Rel.*, Vol. 53, No. 3, 2004, pp. 319–328.

[24]Takaesu, K. and Yoshida, T., "Construction of a Fault-tolerant Voter for N-modular Redundancy," *Electron. Comm. Japan, Part II*, Vol. J84-D-I, No. 4, 2001, pp. 378–388.

[25]Kallenberg, O., *Foundations of Modern Probability*, Springer, New York, NY, 1997.

[26]Burke, C. J. and Rosenblatt, M., "A Markovian Function of a Markov Chain," *Ann. Math. Stat.*, Vol. 29, No. 4, 1958, pp. 1112–1122.

[27]White, L. B., Mahony, R., and Brushe, G. D., "Lumpable Hidden Markov Models – Model Reduction and Reduced Complexity Filtering," *IEEE Trans. Automat. Contr.*, Vol. 45, No. 12, 2000, pp. 2297–2306.

[28]Tejada, A., González, O. R., and Gray, W. S., "On the Markov Property for Nonlinear Discrete-time Systems with Markovian Inputs," *Proc. 2006 American Control Conf.*, Minneapolis, MN, 2004, pp. 899–904.

[29]Zhang, H., Gray, W. S., and González, O. R., "Performance Analysis of Digital Flight Control Systems With Rollback Error Recovery Subject to Simulated Neutron-induced Upsets," *IEEE Trans. Control Syst. Technol.*, Vol. 16, No. 1, 2008, pp. 46–59.

[30]Friedland, B., *Control System Design, An Introduction to State-space Methods*, McGraw-Hill, New York, NY, 1986.