

Fault Coverage Modeling in Nonlinear Dynamical Systems[★]

Matthias A. Müller^a, Alejandro D. Domínguez-García^b

^a*Institute for Systems Theory and Automatic Control, University of Stuttgart, 70550 Stuttgart, Germany*

^b*Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA*

Abstract

In this paper, we propose an approach for modeling fault coverage in nonlinear dynamical systems. Fault coverage gives a measure of the likelihood that a system will be able to recover after a fault occurrence. In our setup, the system dynamics are described by a standard state-space model. The system input (disturbance) is considered to be unknown but bounded at all times. Before any fault occurrence, the vector field governing the system dynamics is such that, for any possible input signal, the corresponding system reach set is contained in some region of the state space defined by the system performance requirements. When a fault occurs, the vector field that governs the system dynamics might be altered. Fault coverage is defined as the probability that, given a fault occurred, the system trajectories remain, at all times, within the region of the state-space defined by the performance requirements. Input-to-State Stability (ISS) concepts are used to compute estimates of the proposed coverage model. Several examples are discussed in order to illustrate the proposed modeling approach.

Key words: Fault Coverage, Fault Tolerance, Nonlinear systems, Invariant sets, Input-to-state stability (ISS).

1 Introduction

Fault tolerance can be defined as the ability of a system to adapt and compensate in a systematic way to random component faults, and keep delivering completely or partially the functionality for which it was designed [11]. Fault tolerance is paramount in control system design for safety- and mission-critical applications. For a specific fault, fault tolerance can be measured through the notion of *fault coverage*, which can be defined as the conditional probability that, given a fault that alters the system structure occurs, the system is able to recover and keeps functioning. The notion of fault coverage was first introduced in the field of fault-tolerant computing (see, e.g., [1,3]), and there is an extensive literature on fault coverage modeling in this field. In this regard, most fault coverage models proposed in fault-tolerant computing are developed using a probabilistic characterization of the fault mechanisms and recovery process; in particular, Markovian models are commonly used (see, e.g., [6,13] and the references therein). For example, in

[18], a continuous-time Markov chain is utilized, where the states of the chain represent the possible outcomes—the system fails or recovers—after the fault occurrence. Then, fault coverage is obtained by computing the probability that the system is in a state of the chain that results in recovery from the fault. For a detailed discussion on fault coverage modeling in fault-tolerant computing, the reader is referred to [6,13].

In contrast with the approaches used in fault-tolerant computing, this paper focuses on fault coverage modeling in nonlinear dynamical systems that can be described by the standard state-space representation used in control theory, $\dot{x} = f(x, u)$, with state x and input u . Throughout this paper, we assume that the system input u is unknown but bounded, which could correspond to an external disturbance or to some uncertainty in the system operation, e.g., uncertainties in the load of an electric power system [15]. The vector field f is defined by the components constituting the system and how these are interconnected. When a fault occurs, the vector field f is altered, resulting in a new vector field \hat{f} . The performance requirements the system has to fulfill are modeled as a certain region in the system state-space to which the system state x is confined. Fault coverage is then defined as the (conditional) probability that the system state x remains inside the region specified by the performance requirements for all times. The goal of the paper is to provide a method for obtaining fault coverage estimates that can be analytically formulated using the system state-space representation.

[★] This paper was not presented at any IFAC meeting. The work of M. A. Müller was supported by the Fulbright Commission. The work of A. D. Domínguez-García was supported in part by the National Science Foundation (NSF) under Career Award ECCS-CAR-0954420.

Email addresses:

matthias.mueller@ist.uni-stuttgart.de
(Matthias A. Müller), aledan@ILLINOIS.EDU
(Alejandro D. Domínguez-García).

The definition of fault coverage adopted in this paper is relevant in various contexts. In a fail-safe application, it might be the case that violating the given performance requirement bounds might be catastrophic, e.g., an aircraft going beyond a critical angle of attack might cause the aircraft to stall. In such a case, it is desirable to guarantee that the system never violates the performance requirements, i.e., to ensure that fault coverage is equal to one independently of the input distribution. However, there might be other applications where performance requirements are “soft” constraints and violating them may cause the system to operate in some degraded mode without catastrophic consequences. In this case, assuming some knowledge of the input distribution, if fault coverage is smaller than one, it does not mean that the system cannot survive to a particular fault, but that after the fault, it might deliver its functionality in a degraded fashion (without necessarily failing catastrophically). For example, in an AC power system, performance requirements impose constraints on frequency deviations, and it is always desirable to keep the frequency within some acceptable range, e.g., 59.6-60.4 Hz. However, temporarily violating this requirement does not mean the system fails catastrophically [15].

Modeling faults as a change in the vector field f is common when designing fault-tolerant controllers (see, e.g., [7,12]), and fault detection filters (see, e.g., [19]). The purpose of fault-tolerant controllers is to ensure stability of the considered system despite the possible occurrence of certain faults. Different methods in achieving this have been proposed in the literature, both for linear and nonlinear systems (see, e.g., [14,2,7,12,21] and the references therein). Considering this, the work in this paper can be seen as a method to quantify the effectiveness of such fault-tolerant controllers, i.e., to give a quantitative expression for the probability that the closed-loop system meets the performance requirements for all times, including the transient phase after the fault occurrence, even if additional uncertainties/disturbances are present.

Fault coverage in the context of fault-tolerant control has also been considered in [20], where a two-level architecture consisting of the considered dynamical system at the lower level and a discrete state Markov process at the upper level, representing the set of possible failures, was used. There, the probabilistic nature of fault coverage stems from a possible uncertainty in the parameters describing a certain fault, and the coverage is intended for decision making in order to prevent the system from entering a state corresponding to a permanent system failure. The framework considered in this paper is significantly different as we explicitly make use of the system dynamics, and utilize reachability analysis techniques to compute fault coverage estimates. In particular, in our framework the probabilistic nature of fault coverage is due to the distribution of the system state at the time of fault occurrence corresponding to the distribution of the input, and not due to parameter uncertainties corresponding to fault occurrences. Our framework provides an analytically tractable method for computing fault coverage estimates, which in turn can directly be incorporated when designing fault-tolerant controllers.

The idea of obtaining fault coverage estimates via the

state space representation of the considered system, which is used in this paper, was already employed in [5] for linear system dynamics. However, the techniques used for computing fault coverage estimates in the nonlinear case are substantially different from the linear case. While ellipsoidal-based reachability analysis techniques are used in [5], in this paper, we use input-to-state stability (ISS) notions.

The structure of this paper is as follows. In Section 2, the system model is presented. This section also provides some background on Input-to-State Stability (ISS) notions, which are key in the development of the proposed fault coverage model. In Section 3, the formal definition of the proposed fault coverage model is given. Section 4 provides analytically tractable methods to compute estimates of the proposed fault coverage model. Section 5 presents several examples that illustrate the ideas developed in the previous sections. Concluding remarks are presented in Section 6.

2 Preliminaries

In this section, we introduce the dynamical system model used throughout the paper, and specify the performance requirements the system is supposed to fulfill. Namely, we assume that our system is described by a state-space representation with an unknown-but-bounded input, and the performance requirements constrain the system trajectories to a region of the state space defined by a symmetric polytope.

2.1 Fault-Free System Dynamics

Let the dynamics of a system operating with no faults be represented by

$$\begin{aligned} \dot{x}(t) &= f(x(t), u(t)), & x(0) &= x_0, \\ u(t) &\in B_u = \{u : |u| \leq u_{max}\}, \end{aligned} \quad (1)$$

where the state $x \in \mathbb{R}^n$, the input $u \in \mathbb{R}^m$, $x_0 \in \mathbb{R}^n$ and $u_{max} \geq 0$. The third equation implies that the input signal u (measurable and locally bounded) is contained in a ball with radius u_{max} for all t . Assume that $f(\cdot, \cdot)$ is locally Lipschitz and the unforced system $\dot{x} = f(x, 0)$ has an asymptotically stable equilibrium point at the origin. We assume that the system (1) is forward complete, i.e., the solution $x(t)$ exists for all $t \geq 0$ and it will be contained in the reachable set $\mathcal{R}(t)$.

2.1.1 Performance Requirements

If the system is properly designed, it must meet some performance requirements. These requirements constrain the state-vector x to some region of the state-space Φ . We assume that Φ is given by a symmetric polytope, defined by

$$\Phi = \{x : |\pi_i^T x| \leq 1, i = 1, 2, \dots, p\}, \quad (2)$$

with $\pi_i \in \mathbb{R}^n$. Then, for every $u(\cdot)$ with $u(t) \in B_u$ for all $t \geq 0$, in order for the system to deliver its intended function, it has to hold that $x(t) \in \Phi$ for all $t \geq 0$, i.e., $\mathcal{R}(t) \subseteq \Phi$ for all $t \geq 0$.

Remark 1 *The subsequent analysis also works if more general constraints than those in the form of symmetric polytopes are considered. However, in this case the*

computation of fault coverage might become more complex or even computationally intractable. Furthermore, performance requirements resulting in polytopical state constraints are fairly general and include many practical problems. \square

2.2 System Dynamics After a Fault

Let T be a random variable representing the time to fault occurrence. This fault alters the vector field f in (1), resulting in a new vector field \hat{f} . Let τ be a realization of T . Then, after a fault, the system state space representation is given by

$$\begin{aligned} \dot{x}(\hat{t}) &= \hat{f}(x(\hat{t}), u(\hat{t})), & x(\hat{t} = 0) &= x(t = \tau) \in \mathcal{R}(\tau), \\ u(\hat{t}) &\in B_u = \{u : |u| \leq u_{max}\}, \end{aligned} \quad (3)$$

where $\hat{t} = t - \tau$ and $\mathcal{R}(\tau)$ is the reach set for system (1) at the time the fault occurs.

2.2.1 Performance Requirements

In fault-free conditions, the performance requirements constrain the system trajectories to some region of the state space Φ . After a fault, the performance requirements imposed on the system might be different (less stringent) than those requirements imposed when the system is operating with no faults, corresponding to a partial functionality after a fault. Therefore, after a fault occurrence, the system trajectories should be constrained to some other region of the state-space, denoted by $\hat{\Phi}$, and defined by

$$\hat{\Phi} = \{x : |\hat{\pi}_i^T x| \leq 1, \forall i = 1, 2, \dots, p\}. \quad (4)$$

2.3 Reachability Bounds

For most nonlinear systems, the exact shape of the reach set $\mathcal{R}(t)$ is impossible to compute and we thus need an upper-bounding approximation. In order to obtain an upper bound for $\mathcal{R}(t)$, we employ the concept of input-to-state stability (ISS). ISS was introduced in [16], and has been extensively studied and applied in systems and control theory in recent years [17]. Loosely speaking, ISS means that bounded inputs lead to bounded system states, and the possible magnitude of the system states scales with the magnitude of the inputs. The precise definition of ISS is given next.

Definition 1 *The system (1) is input-to-state stable (ISS) [16], if there exist functions $\gamma \in \mathcal{K}^1$ and $\beta \in \mathcal{KL}^2$ such that for all $x_0 \in \mathbb{R}^n$ and each input $u(\cdot)$, the solution $x(t)$ of (1) satisfies*

$$|x(t)| \leq \max(\beta(|x_0|, t), \gamma(\|u\|_{[0,t]})), \quad (5)$$

for all $t \geq 0$, where $\|\cdot\|_{[0,t]}$ denotes the supremum norm on the interval $[0, t]$. \square

¹ A function $\alpha: [0, \infty) \rightarrow [0, \infty)$ is of class \mathcal{K} if α is continuous and strictly increasing, and $\alpha(0) = 0$. If α is also unbounded, it is of class \mathcal{K}_∞ .

² A function $\beta: [0, \infty) \times [0, \infty) \rightarrow [0, \infty)$ is of class \mathcal{KL} if $\beta(\cdot, t)$ is of class \mathcal{K} for each fixed $t \geq 0$, and $\beta(r, t)$ decreases to 0 as $t \rightarrow \infty$ for each fixed $r \geq 0$.

The following theorem gives a Lyapunov-like characterization of the input-to-state stability property.

Theorem 1 [10] *Suppose there exist functions $\alpha_1, \alpha_2 \in \mathcal{K}_\infty$, $\rho \in \mathcal{K}$, a continuously differentiable function $V: \mathbb{R}^n \rightarrow \mathbb{R}$ and a continuous, positive definite function $W: \mathbb{R}^n \rightarrow \mathbb{R}$ such that for all $x \in \mathbb{R}^n$*

$$\alpha_1(|x|) \leq V(x) \leq \alpha_2(|x|), \quad (6)$$

$$|x| \geq \rho(|u|) \Rightarrow \frac{\partial V}{\partial x} f(x, u) \leq -W(x). \quad (7)$$

Then the system (1) is ISS with

$$\gamma = \alpha_1^{-1} \circ \alpha_2 \circ \rho. \quad (8)$$

Assumption (6) implies that the ISS-Lyapunov function V is positive definite ($V(0) = 0$ and $V(x) > 0$ for $x \neq 0$) and radially unbounded ($V(x) \rightarrow \infty$ as $|x| \rightarrow \infty$), and assumption (7) ensures that the time derivative of V along trajectories of (1) is negative if the system state at time t , $x(t)$, lies outside a ball of radius $\rho(|u(t)|)$. A proof of Theorem 1 can be found in [10].

If (7) is satisfied not for arbitrarily large x , but only for $x \in B_x$, where B_x is some subset of \mathbb{R}^n including the origin, which is the case if the unforced system is only locally, but not globally, asymptotically stable, then we have to restrict u to be contained in some set B_u such that $\gamma(\|u\|_{[0,t]}) \in B_x$, and thus also x is contained in B_x . The system is then locally ISS.

Assume that after the fault, the system (3) is still ISS. However, the functions $\alpha_1, \alpha_2, \rho, V, W$ and thus also β and γ might be different than before the fault. Theorem 1 is thus valid with some functions $\hat{\alpha}_1, \hat{\alpha}_2, \hat{\rho}, \hat{V}, \hat{W}, \hat{\beta}, \hat{\gamma}$ instead of $\alpha_1, \alpha_2, \rho, V, W, \beta, \gamma$.

Remark 2 *Assuming that the faulty system is still ISS might not be the most general setup; however, this holds true for a variety of practically relevant system classes and typical faults, e.g., line outages in electric power systems as illustrated in the example of Section Section 5 (see, e.g., [15] for a more detailed discussion on power system dynamic models). Furthermore, the framework in this paper can be adapted to the situation where the system is not ISS immediately after the fault, but there is a fault-tolerant controller that reconfigures it so as to make it ISS fast enough. \square*

Now, if the system (1) is ISS and an ISS-Lyapunov function is known, the reach set $\mathcal{R}(t)$ can be bounded from above by some set $\Omega(t)$, given by (5) and (8) respectively:

$$\begin{aligned} \Omega(t) &:= \{x : |x| \leq \max(\beta(|x_0|, t), \gamma(\|u\|_{[0,t]}))\} \\ &= \{x : |x| \leq \max(\beta(|x_0|, t), \alpha_1^{-1}(\alpha_2(\rho(\|u\|_{[0,t]})))\}. \end{aligned} \quad (9)$$

After the transient phase, i.e., for $t \gg 0$, the solution $x(t)$ will then be contained in a set \mathcal{E} , defined by

$$\begin{aligned} \mathcal{E} &:= \{x : |x| \leq \gamma(\sup_{u \in B_u} |u|)\} \\ &= \{x : |x| \leq \alpha_1^{-1}(\alpha_2(\rho(u_{max})))\}. \end{aligned} \quad (10)$$

Furthermore, if x_0 is small enough, i.e., x_0 is contained in some set $B_{x_0} := \{x : V(x) \leq \alpha_2(\rho(u_{max}))\}$, then the solution $x(t)$ will be contained in the set \mathcal{E} for all $t \geq 0$.

Remark 3 Calculating the gain γ via (8) might in general be quite conservative. To overcome this problem, there are methods available in the literature (see, e.g., [9]) to compute a tight value of γ . However, these computations rely on dynamic programming which may result in a heavy computational burden. \square

Remark 4 The estimate \mathcal{E} obtained via (10) might be conservative even if the gain γ is computed optimally, as the estimate \mathcal{E} constitutes a ball regardless of the actual shape of the reach set. Nevertheless, this conservatism can be reduced according to the following considerations. Namely, if \dot{V} can be ensured to be negative outside a certain set \mathcal{B} (not necessarily a ball as in Theorem 1), then the set

$$\mathcal{E}' := \{x : V(x) \leq V_{min}\}, \quad (11)$$

where V_{min} is the smallest possible value such that \mathcal{B} is completely contained in \mathcal{E}' , is invariant with respect to (1), as \dot{V} is negative on all its boundary, and therefore it is a better approximation for the reach set $\mathcal{R}(t)$ (for $t \gg 0$) than \mathcal{E} . Furthermore, if $x_0 \in \mathcal{E}'$, then the solution $x(t)$ will be contained in the set \mathcal{E}' for all $t \geq 0$. However, the shape of the sublevel sets of the ISS-Lyapunov function $V(x)$, and thus also the shape of \mathcal{E}' , might be complicated; thus it might be much easier to work with \mathcal{E} . \square

3 Fault Coverage Definition

In this section, we define the system property of interest—fault coverage. To this end, it is necessary to define the notions of *system failure* and *system recovery*. Note that we adopt the definitions introduced in [5].

Let the dynamics of a system after a first fault be given by (3). Let $\hat{\Phi}$ be the region of the state space defined by the dynamic performance requirements the system must meet after the fault.

Definition 2 *System Failure* [5]. The system fails to deliver its function, if, at the time the fault occurs, the system state variables are such that, for some $\hat{t} > 0$ (with $\hat{t} = t - \tau$), the resulting system trajectory exits the region of the state space defined by $\hat{\Phi}$. \square

Definition 3 *System recovery* [5]. The system survives and recovers from a fault if, at the time the fault occurs, the system state variables are such that the resulting system trajectory remains at all times in the region of the state space defined by $\hat{\Phi}$. \square

Let $\mathcal{R}(\tau)$ be the system reach set at the time of fault occurrence τ . Then, the system survives a fault whenever the state variables at the time of fault occurrence are contained in some set $\hat{\Theta}(\tau) \subseteq \mathcal{R}(\tau) \cap \hat{\Phi}$, such that if $x(t = \tau) = x(\hat{t} = 0) \in \hat{\Theta}(\tau)$, then $x(\hat{t}) \in \hat{\Phi}$ for all $\hat{t} > 0$. This means that if the state x is contained in the set $\hat{\Theta}$ at the time of fault occurrence, then the performance requirements will also be met after the fault. Hence, if $\hat{\Theta}(\tau) \neq \emptyset$, then the system is guaranteed to survive the fault with a probability greater than zero. The definition of fault coverage can now be given as follows.

Definition 4 *Fault Coverage* [5]. Let T be a random variable representing the time to a first fault. Let $X(T)$ be

the random variable representing the system state variables at the time of fault occurrence. Then, for any $t > 0$, fault coverage is defined by

$$C(t) = Pr\{X(T) \in \hat{\Theta}(T) | T < t\}, \quad (12)$$

where $\hat{\Theta}(\tau)$ is the largest set contained in $\mathcal{R}(\tau) \cap \hat{\Phi}$ such that if $x(t = \tau) = x(\hat{t} = 0) \in \hat{\Theta}(\tau)$, then $x(\hat{t}) \in \hat{\Phi}$ for all $\hat{t} > 0$. \square

A case of special interest in the context of fail-safe analysis is when $C = 1$, i.e., the system is guaranteed to survive a fault and recover with probability 1. In Sections 4.1–4.3, we develop a computationally tractable method for computing an estimate of fault coverage defined via (12), before considering the case of fail-safe analysis in more detail in Section 4.4.

4 Fault Coverage Estimation

Several issues make the exact computation of fault coverage as defined in (12) difficult. Namely, the reach set $\mathcal{R}(T)$ as well as the probability distribution of the state variables $X(T)$ over this set is needed. Furthermore, it is necessary to obtain the shape of the set $\hat{\Theta}(\tau)$. Moreover, the probability distribution of the time to fault occurrence T has to be known. Finally, in general, it is not easy to capture the dependence of the distribution of $X(T)$ and the shape of the set $\hat{\Theta}(T)$ on the distribution of T . In the following, we will address each of these issues, introducing simplifying assumptions so as to provide an analytical method to calculate estimates of (12). Furthermore, in Section 4.4 we discuss the situation if no further information about the distribution of the input u is available.

4.1 Time-Scale Separation of System Dynamics and Fault Occurrence [5]

Fault coverage estimation can be simplified if we assume that the time constants associated with the system dynamics are much smaller than the time constants associated with fault occurrences. This is known as behavioral decomposition [5], and it is a quite reasonable assumption for a variety of system classes, such as aerospace systems, automotive systems, or power systems. Namely, for such systems, the system dynamics time constants are in the order of seconds, while typical fault rates for reasonably reliable systems are in the order of $10^{-5} - 10^{-9}$ /h. While in the general case, fault coverage defined as in (12) also depends on the initial condition x_0 of the system, this dependence is removed if behavioral decomposition holds.

Let $f_{X|T}(x|\tau)$ represent the (conditional) probability density function of $X(T)$. As was shown in [5], (12) can be simplified using the behavioral decomposition assumption, and one obtains

$$C \approx \int_{x(\tau) \in \hat{\Theta}_{ss}} f_{X|T}(x|\tau) dx, \quad (13)$$

where $\hat{\Theta}_{ss}$ is the steady-state value of $\hat{\Theta}(\tau)$. In the following two sections, we will derive a procedure how the domain of integration $\hat{\Theta}_{ss}$ and the probability density function $f_{X|T}$ can be computed, such that fault coverage estimates can be obtained via (13).

4.2 Approximation of the Domain of Integration $\hat{\Theta}_{ss}$

In order to compute fault coverage estimates according to (13), we need the set $\hat{\Theta}_{ss}$. As stated before, $\hat{\Theta}_{ss}$ is the steady-state value of the largest set $\hat{\Theta}(\tau) \subseteq \mathcal{R}(\tau) \cap \hat{\Phi}$ such that if $x(\tau) \in \hat{\Theta}_{ss}$ at the time of fault occurrence, then $x(\hat{t}) \in \hat{\Phi}$ for all $\hat{t} > 0$. Since $\mathcal{R}(\tau)$ is difficult to compute, so is $\hat{\Theta}(\tau)$; thus we use the approximation \mathcal{E} (or \mathcal{E}' , respectively) of the steady-state value of $\mathcal{R}(t)$ defined in (10) (or (11), respectively) and compute, instead of $\hat{\Theta}_{ss}$, the largest (in some sense) set $\hat{\mathcal{E}} \subseteq \mathcal{E} \cap \hat{\Phi}$ such that if $x(\tau) \in \hat{\mathcal{E}}$ at the time of fault occurrence, then $x(\hat{t}) \in \hat{\Phi}$ for all $\hat{t} > 0$. This can be accomplished by first obtaining the largest (in some sense) set $\hat{\mathcal{X}}$ which is invariant with respect to the system dynamics (3) after the fault occurrence, and which is completely contained in $\hat{\Phi}$; and then by obtaining $\hat{\mathcal{E}}$ as the intersection of \mathcal{E} and $\hat{\mathcal{X}}$. The previously sketched procedure for obtaining the set $\hat{\mathcal{E}}$ is discussed next in detail.

We compute the set $\hat{\mathcal{X}}$ as the largest sublevel set of the ISS-Lyapunov function $\hat{V}(x)$, such that $\hat{\mathcal{X}}$ is both invariant with respect to (3) and contained in $\hat{\Phi}$. Let $a > 0$. Consider the sublevel set $\mathcal{P}_a = \{x : \hat{V}(x) \leq a\}$, whose boundary $\partial\mathcal{P}_a$ is the level set $\partial\mathcal{P}_a = \{x : \hat{V}(x) = a\}$. According to (6), we get $|x| \geq \hat{\alpha}_2^{-1}(a)$ for all $x \in \partial\mathcal{P}_a$. Hence if $\hat{\alpha}_2^{-1}(a) \geq \hat{\rho}(u_{max})$, then $|x| \geq \hat{\rho}(u_{max})$ for all $x \in \partial\mathcal{P}_a$ and thus, according to (7), $\frac{d}{dt}\hat{V}$ is negative for all $x \in \partial\mathcal{P}_a$. As $\frac{d}{dt}\hat{V}$ is negative on all its boundary, the set \mathcal{P}_a is invariant with respect to (3). Hence the set $\hat{\mathcal{X}}$ is given by $\hat{\mathcal{X}} = \mathcal{P}_{a_{max}} = \{x : \hat{V}(x) \leq a_{max}\}$, where a_{max} is obtained with the following optimization problem:

$$\text{maximize} \quad a \quad (14)$$

$$\text{subject to} \quad \mathcal{P}_a \in \hat{\Phi} \quad (15)$$

$$a \geq \hat{\alpha}_2(\hat{\rho}(u_{max})) \quad (16)$$

As stated above, the second constraint (16) ensures that the set $\hat{\mathcal{X}}$ is invariant with respect to (3). Furthermore, the complexity of the optimization problem is determined by the shape of the sublevel sets of \hat{V} , according to (15). If, for example, \hat{V} is a quadratic ISS Lyapunov function, i.e., $\hat{V}(x) = x'\hat{\Xi}^{-1}x$, where $\hat{\Xi}$ is a positive definite symmetric matrix, then the level set $\hat{V}(x) = a_{max}$ is an ellipsoid and the requirement (15) that $\hat{\mathcal{X}}$ has to be contained in $\hat{\Phi}$ translates into [4]

$$a_{max}\hat{\pi}_i^T\hat{\Xi}\hat{\pi}_i \leq 1, \quad \forall i = 1, 2, \dots, p,$$

which results in a tractable convex optimization problem [4]. If the shape of the level sets is more complicated and thus the constraint (15) is not tractable anymore, a more conservative but very easily tractable constraint can be used instead to ensure that the set $\hat{\mathcal{X}}$ is contained in $\hat{\Phi}$. Namely, according to (6), $|x| \leq \hat{\alpha}_1^{-1}(a_{max})$ for all $x \in \partial\hat{\mathcal{X}}$. Thus, if we ensure that the ball $B_{out} =$

$\{x : |x| \leq \hat{\alpha}_1^{-1}(a_{max})\}$ is contained in $\hat{\Phi}$, then also $\hat{\mathcal{X}}$, which lies inside B_{out} , is contained in $\hat{\Phi}$. This can easily be ensured by requiring that $|\hat{\pi}_i| \leq 1/\hat{\alpha}_1^{-1}(a_{max})$ for all $i = 1, 2, \dots, \hat{p}$, or equivalently,

$$a_{max} \leq \hat{\alpha}_1(1/|\hat{\pi}_i|). \quad (17)$$

Then, according to the definition of the performance requirement region (4), $\hat{\mathcal{X}}$ lies inside the symmetric polytope, as $|\hat{\pi}_i^T x| \leq |\hat{\pi}_i||x| \leq |\hat{\pi}_i|\hat{\alpha}_1^{-1}(a_{max}) \leq |\hat{\pi}_i|\hat{\alpha}_1^{-1}(\hat{\alpha}_1(1/|\hat{\pi}_i|)) = 1$. The tightest constraint on a_{max} is the one where $|\hat{\pi}_i|$ is maximal, i.e., (17) reduces to

$$a_{max} \leq \hat{\alpha}_1\left(\frac{1}{\max_{i \in \{1, 2, \dots, \hat{p}\}} |\hat{\pi}_i|}\right).$$

With this, the optimization problem (14) - (16) reduces to checking if

$$\hat{\alpha}_1\left(\frac{1}{\max_{i \in \{1, 2, \dots, \hat{p}\}} |\hat{\pi}_i|}\right) < \hat{\alpha}_2(\hat{\rho}(w_{max})),$$

in which case $\hat{\mathcal{X}}$ is an empty set and fault coverage is zero, or if

$$\hat{\alpha}_1\left(\frac{1}{\max_{i \in \{1, 2, \dots, \hat{p}\}} |\hat{\pi}_i|}\right) \geq \hat{\alpha}_2(\hat{\rho}(w_{max})),$$

in which case the set $\hat{\mathcal{X}}$ is given by

$$\hat{\mathcal{X}} = \{x : \hat{V}(x) \leq \hat{\alpha}_1\left(\frac{1}{\max_{i \in \{1, 2, \dots, \hat{p}\}} |\hat{\pi}_i|}\right)\}. \quad (18)$$

As described above, in order to obtain $\hat{\mathcal{E}}$, we have to calculate the intersection of \mathcal{E} and $\hat{\mathcal{X}}$, where according to (10), \mathcal{E} is a n-dimensional sphere and $\hat{\mathcal{X}}$ is given by a sublevel set of \hat{V} . As above, if those sublevel sets have a complicated shape, it might not be possible to easily calculate the intersection of \mathcal{E} and $\hat{\mathcal{X}}$. But again, as above, a more conservative but easily computable estimate for $\hat{\mathcal{E}}$ can be obtained. Namely, consider the biggest sphere which is completely contained in $\hat{\mathcal{X}}$, which according to (6) is given by $B_{max} = \{x : |x| \leq \hat{\alpha}_2^{-1}(a_{max})\}$. Then, $\hat{\mathcal{E}}$ is just the smaller one of the two spheres \mathcal{E} and B_{max} , namely

$$\hat{\mathcal{E}} = \{x : |x| \leq \min\{\hat{\alpha}_1^{-1}(\hat{\alpha}_2(\hat{\rho}(u_{max}))), \hat{\alpha}_2^{-1}(a_{max})\}\}. \quad (19)$$

If the approximation for $\hat{\mathcal{X}}$ given in (18) is used, (19) yields $\hat{\mathcal{E}} = \{x : |x| \leq \min\{\rho_1, \rho_2\}\}$ with $\rho_1 = \hat{\alpha}_1^{-1}(\hat{\alpha}_2(\hat{\rho}(u_{max})))$ and $\rho_2 = \hat{\alpha}_2^{-1}(\hat{\alpha}_1(\frac{1}{\max_i |\hat{\pi}_i|}))$.

4.3 Determination of State Variables Distribution

Obtaining the probability density function of $X(T)$, which we denote by $f_{X|T}(x|t)$, might be a challenging task, and sometimes even impossible, task. In general, $f_{X|T}(x|t)$ depends on the time structure of the input u and its distribution over the set B_u . However, in many systems, there is only partial information about

the time structure of the system input, and thus it is not possible to completely determine $f_{X|T}(x|t)$. Nevertheless, under the additional assumption that u is quasistatic with respect to the system dynamics (1), i.e., the timeframe for changes in the value of u is much larger than the time constants of the system dynamics (1) and thus the system is in steady-state most of the time, we can compute the steady-state probability density function of $X(T)$, as shown below. There are many practical systems in which this condition holds. For example, in the power system considered in Section 5.2, the voltage at the infinite bus, V_∞ , can be assumed to be constant for a certain period of time, before it changes to a different value and remains constant until another change occurs. The steady-state of the system (1) is given by $0 = f(x, u)$. Suppose that we can solve this for $x = g(u)$, where g is a continuously differentiable function $g: \mathbb{R}^m \rightarrow \mathbb{R}^n$ with full rank Jacobian, i.e., the rank r of the Jacobian is given by $r = \min(m, n)$. If the probability density function of U , which we denote by $f_U(u)$, is known, for the case of a square system, i.e., $m = n$, the (steady-state) probability density function of X can then be computed by [8]

$$f_{X|T}(x|\tau) = \begin{cases} \frac{f_U(u)}{|\frac{dx}{du}|} = f_U(u) \left| \frac{du}{dx} \right| & \text{if } x \in \mathcal{R}_{ss}, \\ 0 & \text{else} \end{cases} \quad (20)$$

where $|\frac{dx}{du}|$ denotes the determinant of the Jacobian $\frac{dx}{du}$, and \mathcal{R}_{ss} is the set of reachable equilibrium points of the system (1), given by $\mathcal{R}_{ss} = \{x : x = g(u), u \in B_u\}$. Note that $\mathcal{R}_{ss} \subseteq \mathcal{E}' \subseteq \mathcal{E}$.

If the system is non-square, we have to distinguish two different cases. If $m > n$, then the (steady-state) probability density function of X can be obtained by adding $m - n$ new coordinates x_{n+1}, \dots, x_m and extending g such that $g: \mathbb{R}^m \rightarrow \mathbb{R}^m$ has a non-singular Jacobian $|\frac{dg}{du}|$. Then we can compute $f_{X|T}(x|t)$ by first calculating the m -dimensional density (according to (20)), and then integrating out the coordinates x_{n+1}, \dots, x_m .

If $m < n$, then, as the Jacobian of g has rank m , by the inverse function theorem it is possible to take the first m coordinates³ x_1, \dots, x_m and express the remaining $n - m$ coordinates x_{m+1}, \dots, x_n through x_1, \dots, x_m , i.e., there exists a function $h: \mathbb{R}^m \rightarrow \mathbb{R}^{n-m}$ such that

$$[x_{m+1}, \dots, x_n] = h(x_1, \dots, x_m). \quad (21)$$

The (steady-state) probability density function of X is then given by the m -dimensional density for x_1, \dots, x_m , which can be obtained from (20), and \mathcal{R}_{ss} thus lies on an m -dimensional manifold given by (21).

4.4 Generalization to fail-safe analysis

Summarizing the above, we provided an analytically tractable method to compute fault coverage estimates for the case where behavioral decomposition holds and the input u is quasi-static with respect to the system

³ Assume without loss of generality that $[g_1, \dots, g_m]^T$ has rank m ; if not, pick m coordinates x_{i_1}, \dots, x_{i_m} such that the Jacobian of $[g_{i_1}, \dots, g_{i_m}]^T$ has rank m .

dynamics, and furthermore the distribution of the input over the set B_u is known. However, in practice often no information on the distribution of the input u over the set B_u is known, but only that it is bounded. In this case it is not possible to obtain the probability density function $f_{X|T}$ needed for calculating fault coverage estimates via (12) and (13), respectively. Nevertheless, the above definition of fault coverage can still be used in the context of a fail-safe analysis, i.e., for checking whether $C = 1$, which means that the system is guaranteed to survive a fault for any given input $u \in B_u$. For the case that both the assumptions of behavioral decomposition as well as u being quasi-static with respect to the system dynamics hold, we obtain that $C = 1$ if $\mathcal{R}_{ss} \subseteq \hat{\mathcal{X}}$. If the assumption of u being quasi-static is dropped, this holds true if $\mathcal{E} \subseteq \hat{\mathcal{X}}$. If furthermore the assumption of behavioral decomposition is removed, we obtain that $C = 1$ if $\Omega(t) \subseteq \hat{\mathcal{X}}$ for all $t \geq 0$, with Ω defined in (9). We will illustrate this fail-safe analysis with the examples in Section 5.

5 Examples

The purpose of this section is to illustrate the proposed method of obtaining fault coverage estimates introduced in the previous sections. We first apply the results to a first-order linear RL circuit, and then show the applicability of our results to a simple (nonlinear electric) power system—the so-called single-machine infinite-bus (SMIB) system [15].

5.1 A First-Order Linear System: RL Circuit

Consider a series RL circuit, with an inductor of value L , a resistor of value R_{eq} that results from connecting in parallel two resistors of values R_1 and R_2 , i.e., $R_{eq} = \frac{R_1 R_2}{R_1 + R_2}$, and a voltage source of value $v(t)$ driving in series with L and R_{eq} . We assume that the voltage source $v(t)$ is unknown, but it is such that $|v(t)| \leq V, \forall t \geq 0$, for some $V > 0$. The maximum currents that resistors R_1 and R_2 can process are $i_{max}^{R_1}$ and $i_{max}^{R_2}$, with $V/R_1 < i_{max}^{R_1}$ and $V/R_2 < i_{max}^{R_2}$, respectively, and once these are reached the resistors fail open. We consider possible faults caused by resistor open-circuit failure.

Before any fault occurrence, the current $i(t)$ is governed by

$$\frac{di(t)}{dt} = -\frac{R_{eq}}{L}i(t) + \frac{1}{L}v(t) =: f(i(t), v(t)), \quad v(t) \in B_v := \{v : |v| \leq V\}, \quad (22)$$

where $R_{eq} = \frac{R_1 R_2}{R_1 + R_2}$. Assume that at some time τ resistor R_1 fails open circuit. This alters the vector field f in such a way that R_{eq} is replaced by R_2 , i.e.,

$$\hat{f}(i(\hat{t}), v(\hat{t})) := -\frac{R_2}{L}i(\hat{t}) + \frac{1}{L}v(\hat{t}), \quad (23)$$

where $\hat{t} = t - \tau$. The maximum current that can flow through the circuit after the fault is limited by $i_{max}^{R_2}$, which is the maximum current allowed through R_2 . Thus, a system failure occurs if for some $\hat{t} > 0$, $i(\hat{t}) \notin \hat{\Phi} := \{i : |i| \leq i_{max}^{R_2}\}$.

Using $V(i) = \frac{1}{2}i^2$ as an ISS-Lyapunov function for the system (22), we obtain the sets \mathcal{E} and \mathcal{E}' , given by (10) and (11), respectively, as

$$\mathcal{E} = \mathcal{E}' = \{i : |i| \leq \rho(V)\} = \{i : |i| \leq V/R_{eq}\}. \quad (24)$$

For the post-fault system (23), we can use the same ISS-Lyapunov function as for the fault-free system and follow the procedure described in Section 4.2 to obtain $\hat{\mathcal{X}} = \{i : \hat{V}(i)\} = \{i : |i| \leq i_{max}^{R_2}\}$ via (18). The set $\hat{\mathcal{E}}$ can now be computed as the intersection of the set \mathcal{E} and $\hat{\mathcal{X}}$. We have to distinguish two cases: If $V/R_{eq} \leq i_{max}^{R_2}$, then $\hat{\mathcal{E}} = \mathcal{E} \subseteq \hat{\mathcal{X}}$, and the fault coverage is $C = 1$. Hence from a fail-safe analysis point of view, the system is guaranteed to survive a failure of resistor R_2 without any further assumptions on the distribution of v . On the other hand, if $V/R_{eq} > i_{max}^{R_2}$, then $\hat{\mathcal{E}} = \hat{\mathcal{X}} \subset \mathcal{E}$, and hence $C < 1$. This means that from a fail-safe analysis point of view, the system is not guaranteed to survive a failure of resistor R_2 for any given input v . If we furthermore assume that the random variable V associated with the input voltage is uniformly distributed over B_v and that the voltage $v(t)$ is quasi-static with respect to the circuit dynamics, then, according to (20), the random variable $I(T)$ associated with the current in the circuit at the time of fault occurrence is uniformly distributed over the set of reachable equilibrium points \mathcal{R}_{ss} , which in this example coincides with the set \mathcal{E} (24). Hence in this case, we obtain that

$$C = \frac{\text{vol}(\hat{\mathcal{E}})}{\text{vol}(\mathcal{E})} = \frac{\text{vol}(\hat{\mathcal{X}})}{\text{vol}(\mathcal{E})} = \frac{R_{eq, R_2}}{V}.$$

This is the same result obtained in [5], where ellipsoidal techniques to approximate the reach set are used.

5.2 A Second-Order Nonlinear System: Single-Machine Connected to an Infinite Bus

The goal of this example is to show the applicability of our results to a simple nonlinear electric power system, but also to illustrate the problem of finding a good ISS-Lyapunov function, i.e., an ISS-Lyapunov function such that the results are not too conservative. In particular, the example shows that better results can be obtained if the shape of the level sets of the ISS-Lyapunov function is known.

The system under consideration is given by an electric power generator connected by two lines in parallel, with complex impedances jX_1 and $2jX_1$, respectively, to an infinite bus with voltage V_∞ . Let δ be the angular position of the rotor in electrical radians, and ω_r be the angular velocity of the rotor in electrical rad/s. Then, the system dynamics can be described by

$$\begin{aligned} \dot{\delta} &= \omega_r - \omega_s \\ \dot{\omega}_r &= -Da\omega_r - ab \sin \delta V_\infty + aT_m + Da\omega_s, \end{aligned} \quad (25)$$

with

$$V_\infty \in \Omega_{V_\infty} := \{V_\infty : |V_\infty - V| \leq kV\},$$

for some $k > 0$ and $V > 0$, and $a = \frac{\omega_s}{2H}$ and $b = \frac{3E_t}{3X_m + 2X_1}$, where D, H, E_t, X_m, ω_s , and T_m are constant parameters [15].

The (stable) equilibrium point of the fault-free system (25) is given by $\omega_r^s = \omega_s$ and $\delta^s = \sin^{-1} T_m/(bV) \in [0, \pi/2]$. Applying the coordinate change $x_1 = \delta - \delta^s$, $x_2 = \omega_r - \omega_r^s$ in order to shift the equilibrium point to the origin and writing $V_\infty = V + \Delta V$, with $|\Delta V| \leq kV$ according to (25), results in the following system:

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -Da x_2 - h(x_1) - ab \sin(x_1 + \delta^s) \Delta V, \end{aligned} \quad (26)$$

where $h(x_1) = abV \sin(x_1 + \delta^s) - aT_m$. Using $V(x) = 0.5x'Px + \int_0^{x_1} h(y)dy$ with

$$P = \begin{bmatrix} \epsilon a^2 D^2 & \epsilon a D \\ \epsilon a D & 1 \end{bmatrix},$$

for some $0 < \epsilon < 1$ as an ISS-Lyapunov function, we can establish that the system (26) is locally ISS in the region $\bar{\mathcal{D}} := \{x : |x_1| \leq \frac{\pi}{2} - \delta^s\}$.

Now suppose that the performance requirements are such that, first, we want to stay in the region $x \in \bar{\mathcal{D}}$ for which we established ISS; i.e., we want to have

$$|x_1| \leq \frac{\pi}{2} - \delta^s. \quad (27)$$

Second, we require that the electrical frequency deviations are smaller than 1 Hz, which is equivalent to requiring that the deviations in the machine speed have to be less than 2π rad/s, i.e.,

$$|x_2| \leq 2\pi. \quad (28)$$

If we work out a numerical example with typical parameter values, it turns out that the estimate for the reach set we obtain from (10) is too conservative. Namely, only for very small deviations in the input voltage, we can satisfy the performance requirements. However, if we use our knowledge of the shape of the level sets of the function $V(x)$, we get a better estimate. Note that these level sets have a shape close to an ellipsoid (squeezed in the x_1 -direction by the integral term in V). As explained in Remark 4 in Section 2.3, we compute \mathcal{E}' according to (11) as a better approximation for the reach set after the transient phase than the set \mathcal{E} we get from (10).

As a specific example, consider a SMIB with the following parameter values: $E_t = 1$ pu, $X_m = 0.2$ pu, $X_1 = 0.1$ pu, $H = 4$ pu, $D = 0.04$ s/rad, $T_m = 1$, $\omega_s = 120\pi$ rad/s, and $V = 1$ pu. Taking $\epsilon = 0.3$, it turns out that if $k \leq 0.085$, i.e., if the deviations in the input voltage V are smaller than 8.5%, then the set \mathcal{E}' is contained in the region specified by the performance requirements (27)–(28). Furthermore, it turns out that the constraint on the deviations in the electrical frequency (28) is much tighter than the constraint on x_1 (27) in order to remain in the ISS region. If k is further increased beyond $k = 0.085$, then we cannot ensure anymore that the second performance requirement is met.

Assume now that a fault in the line with impedance jX_1 will cause this line to open. The system dynamics are then given as in (25), but with b replaced by $\hat{b} = E_t/(X_m + 2X_1)$. Following the same steps as in the fault-free case, one can establish that the faulty system is still

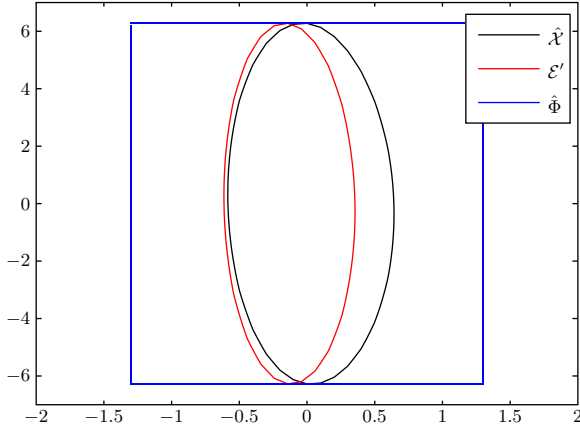


Fig. 1. The sets $\hat{\Phi}$, \mathcal{E}' , and $\hat{\mathcal{X}}$ for the SMIB system.

locally ISS, and the set $\hat{\mathcal{X}}$ can be computed as described in Section 4.2 as the largest invariant sublevel set of \hat{V} which lies in the region specified by the performance requirements, which are assumed to remain the same after the fault. For the given numerical example, the sets $\hat{\mathcal{X}}$, $\hat{\Phi}$ and \mathcal{E}' are depicted in Fig. 1.

To compute the actual values for the fault coverage, we need the distribution of x over the estimate of the reach set \mathcal{E}' . We do not completely know the time structure of the infinite bus voltage V_∞ ; however, as stated above, it is reasonable to assume that the input voltage V_∞ is quasi-static with respect to the system dynamics, and thus we can calculate the steady-state probability density function $f_{X|T}$ as described in Section 4.3. For the given numerical example, it turns out that the set \mathcal{R}_{ss} is contained in $\hat{\mathcal{X}}$, and hence fault coverage is $C = 1$. This means that the system is fail-safe if the input voltage V_∞ is quasi-static with respect to the system dynamics with deviations from its nominal value less than 8.5%.

On the other hand, if the assumption of V_∞ being quasi-static is removed, we cannot guarantee that the SMIB is fail-safe for input voltage deviations up to 8.5%, as \mathcal{E}' is not completely contained in $\hat{\mathcal{X}}$ and hence $C < 1$ as can be seen in Fig. 1. However, one can show that the post-fault system remains within a region such that $|\hat{x}_2| \leq 2.2\pi$, which means that the electrical frequency can be guaranteed to reside in a region slightly larger than specified by the performance requirement (28). This can be interpreted as that the SMIB still delivers its functionality but in a slightly degraded fashion, yet without failing catastrophically, i.e., without ceasing operation.

6 Concluding Remarks

In this paper, we proposed a fault coverage model for nonlinear dynamical systems described by a state space representation, where the inputs are unknown but bounded. An analytically tractable method to compute estimates of this fault coverage model was provided. This method builds on input-to-stability (ISS) notions and generalizes previous results for linear systems that used ellipsoidal-based reachability analysis techniques.

References

- [1] T. Arnold. The concept of coverage and its effect on the reliability model a repairable system. *IEEE Transactions on Computers*, C-22(3):251–254, March 1973.
- [2] C. Bonivento, A. Isidori, L. Marconi, and A. Paoli. Implicit fault-tolerant control: application to induction motors. *Automatica*, 40(3):355 – 371, 2004.
- [3] W. Bouricius, W. Carter, and P. Schneider. Reliability modeling techniques for self-repairing computer systems. In *Proceedings of the 24th National ACM Conference*, New York, NY, 1969. ACM Press.
- [4] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, UK, 2004.
- [5] A. Dominguez-Garcia, J. Kassakian, and J. Schindall. A generalized fault coverage model for linear time-invariant systems. *IEEE Transactions on Reliability*, 58(3):553–567, September 2009.
- [6] J. Dugan and K. Trivedi. Coverage modeling for dependability analysis of fault-tolerant systems. *IEEE Transactions on Computers*, 38(6):775–787, June 1989.
- [7] S. Gao, Z. and Ding. Actuator fault robust estimation and fault-tolerant control for a class of nonlinear descriptor systems. *Automatica*, 43:912–920, May 2007.
- [8] G. Grimmett and D. Stirzaker. *Probability and Random Processes*, 3rd ed. Oxford University Press, Oxford, UK, 2001.
- [9] S. Hunag, M. R. James, D. Nesic, and P. M. Dower. Analysis of input-to-state stability for discrete time nonlinear systems via dynamic programming. *Automatica*, 41(12):2055–2065, 2005.
- [10] H.K. Khalil. *Nonlinear Systems*. Prentice-Hall, Upper Saddle River, NJ, 2002.
- [11] J. Laprie. *Dependability: Basic Concepts and Terminology*. Springer-Verlag, New York, NY, 1991.
- [12] Z. Mao, B. Jiang, and P. Shi. Fault-tolerant control for a class of nonlinear sampled-data systems via a euler approximate observer. *Automatica*, 46:1852–1859, November 2010.
- [13] D. Pradhan, editor. *Fault-Tolerant Computer System Design*. Prentice Hall, New Jersey, NJ, 1995.
- [14] Zhihua Qu, Curtis M. Ihlefeld, Yufang Jin, and Apiwat Saengdeejing. Robust fault-tolerant self-recovering control of nonlinear uncertain systems. *Automatica*, 39(10):1763 – 1771, 2003.
- [15] P. Sauer and A. Pai. *Power System Dynamics and Stability*. Prentice Hall, Upper Saddle River, NJ, 1998.
- [16] E. D. Sontag. Smooth stabilization implies coprime factorization. *IEEE Transactions on Automatic Control*, 34:435–443, 1989.
- [17] E. D. Sontag. Input to state stability: Basic concepts and results. In P. Nistri and G. Stefani, editors, *Nonlinear and Optimal Control Theory*, pages 163–220. Springer-Verlag, Berlin, 2007.
- [18] J. Stiffler and L. Bryant. Care III phase II report – mathematical description. Technical Report NASA-CR-3566, NASA, November 1982.
- [19] A. Willsky. A survey of design methods for failure detection systems. *Automatica*, 12(6):601–611, November 1976.
- [20] N. E. Wu. Coverage in fault-tolerant control. *Automatica*, 40:537–548, April 2004.
- [21] H. Yang, B. Jiang, and M. Staroswiecki. Supervisory fault tolerant control for a class of uncertain nonlinear systems. *Automatica*, 45:2319–2324, October 2009.