An Event-based Approach to Hybrid Systems Diagnosability

Matthew Daigle*† and **Xenofon Koutsoukos**† and **Gautam Biswas**†
*University of California, Santa Cruz, NASA Ames Research Center, Moffett Field, CA, USA
†Institute for Software Integrated Systems, Vanderbilt University, Nashville, TN, USA
mdaigle@arc.nasa.gov

Abstract

Diagnosability is an important issue in the design of diagnostic systems, because it helps identify whether sufficient information is available to distinguish all the faults. Diagnosability of hybrid systems, however, is challenging, because mode transitions may occur during fault isolation. We present an event-based framework for hybrid systems diagnosis based on a qualitative abstraction of measurement deviations from nominal behavior. We derive event-based fault models that describe the possible measurement deviations sequences due to faults, which, coupled with the mode transition structure of the system, are used to automatically synthesize an event-based diagnoser for hybrid systems. We introduce notions of diagnosability for hybrid systems and show how the event-based diagnoser can be used to verify the diagnosability of the system. We apply our diagnosability analysis scheme to a real-world electrical power distribution system.

1 Introduction

Diagnosability relates to the ability of a diagnostic system to obtain unique diagnosis results given a set of observations. Therefore, it is an important property that affects many aspects of the design of diagnostic systems. Based on diagnosability, we can determine at design time if a set of sensors provide sufficient discriminatory evidence, and, if not, what additional sensors may be useful.

Many modern engineering systems are best modeled as hybrid systems, which combine continuous and discrete behaviors in a common framework. Yet, diagnosability analysis of hybrid systems has largely been ignored. The task is complicated, because the effects of faults may change from one mode to another. In discrete-event systems, diagnosability refers to obtaining a sequence of observable events that is unique enough to identify which failure has occurred [Sampath *et al.*, 1995; Zad *et al.*, 2003]. Diagnosability of continuous systems has also been well-studied [Travé-Massuyès *et al.*, 2006], and can be seen in much the same way, if fault signatures are viewed as events [Cordier *et al.*, 2006; Daigle *et al.*, 2007a; Meseguer *et al.*, 2008]. Diagnosability of hybrid systems is studied in [Benedetto *et al.*, 2007], but is defined only as the ability to detect faults, and not to obtain unique isolation results. Hybrid systems diagnosability in the analytic redundancy relations framework is described in [Bayoudh *et al.*, 2006], and accounts for the changes in fault signatures due to mode changes.

We adopt an event-based approach to hybrid systems diagnosability, where faults are viewed as unobservable events. Measurement deviations (i.e., fault signatures) and controlled mode changes form the set of observable events. As in [Sampath *et al.*, 1995; Cordier *et al.*, 2006], we say a system is diagnosable if the sequence of observable events after fault occurrence uniquely isolates the fault. Due to mode changes, diagnosability of hybrid systems is typically harder to achieve than for continuous systems. A hybrid system might be diagnosable within each individual mode, but mode transitions during the fault isolation process may lead to loss of diagnosability because fault effects could get masked. Therefore, we introduce the more practical notion of Q-diagnosability, in which diagnosability can be achieved by blocking or forcing certain controlled mode changes during fault isolation. We design event-based diagnosers, which are then used to verify the diagnosability properties of the system. We apply our diagnosability scheme to a subset of the Advanced Diagnostics and Prognostics Testbed (ADAPT) at NASA Ames, which is a complex electrical power distribution system.

The paper is organized as follows. Section 2 describes the qualitative fault isolation framework. Section 3 presents the event-based fault modeling approach. Section 4 formalizes diagnosability in our framework, and Section 5 describes the design of the event-based diagnoser and how it can be used to verify diagnosability. Section 6 presents the case study. Section 7 concludes the paper.

2 Qualitative Fault Isolation

We consider the problem of single fault diagnosis in hybrid systems. We represent faults as unobservable events, and consider both abrupt parametric faults, modeled as unexpected step changes in system parameter values, and discrete faults, modeled as unexpected changes in system mode. Nominal mode transitions can occur due to known external controller actions, or autonomous behaviors that depend on internal system variables. In this paper, we assume that autonomous mode changes do not occur during fault isolation. Autonomous modes changes and multiple faults

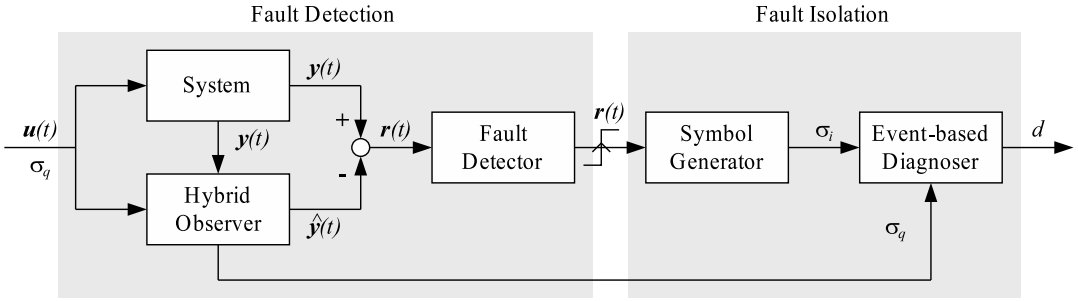


Figure 1: Event-based diagnosis architecture.

can be incorporated in a more complex framework using the techniques presented in [Narasimhan and Biswas, 2007; Daigle, 2008]. This paper does not consider these extensions to focus on the notions of diagnosability for hybrid systems.

The hybrid diagnosis architecture is illustrated in Fig. 1. A hybrid observer, implemented as a switched extended Kalman filter, computes the expected behavior of the plant based on inputs $\mathbf{u}(t)$ and controlled mode change commands σ_q [Narasimhan and Biswas, 2007]. The difference between observed outputs, $\mathbf{y}(t)$, and expected outputs, $\hat{\mathbf{y}}(t)$, defines the residual, $\mathbf{r}(t)$. The fault detector employs a statistical test of significance to robustly determine if the residual is nonzero using a sliding window technique [Biswas *et al.*, 2003]. Measurement deviations from nominal behavior are abstracted via the symbol generator, and the event-based diagnoser uses the sequence of events formed by measurement deviations, σ_i, and controlled mode changes, σ_q, to isolate faults. In the following, we denote the set of modes as $Q = \{q_1, q_2, \ldots, q_r\}$, the set of faults as $F = \{f_1, f_2, \ldots, f_n\}$, and the set of measurements, which are time-varying signals obtained from the available sensors, as $M = \{m_1, m_2, \ldots, m_p\}$.

Measurement deviations are abstracted using qualitative +, −, and 0 values to form *fault signatures* [Mosterman and Biswas, 1999]. Fault signatures represent the immediate change in magnitude and the first nonzero derivative change. They also represent what is termed discrete change behavior, which describes whether the signal went from a nonzero to a zero value (Z), a zero to a nonzero value (N), or had no zero/nonzero value changes (X) [Daigle *et al.*, 2008].

Definition 1 (Fault Signature)**.** A *fault signature* for a fault f and measurement m in mode q is the qualitative magnitude, slope, and discrete change in m caused by the occurrence of f, and is denoted by $\sigma_{f,m,q} \in \Sigma_{f,m,q}$. We denote the set of all fault signatures for fault f and measurements M in mode q as $\Sigma_{f,M,q}$, where $\Sigma_{f,M,q} = \bigcup_{m \in M} \Sigma_{f,m,q}$.

If the fault signature for a fault f and measurement m can be uniquely determined, then $\Sigma_{f,m,q}$ is a singleton. In general, $\sigma_{f,m}$ may not be unique due to ambiguities in the qualitative arithmetic.

In addition to fault signatures, we also capture the temporal order of measurement deviations, termed *relative measurement orderings* [Daigle *et al.*, 2007b], which refer to the intuition that fault effects will manifest in some parts of the system before others. Measurement orderings are based on analysis of the transfer functions from faults to measurements [Daigle *et al.*, 2007b].

Definition 2 (Relative Measurement Ordering)**.** If fault f manifests in measurement m_i before measurement m_j in mode q, then we define a *relative measurement ordering* between m_i and m_j for fault f in q, denoted by $m_i \prec_{f,q} m_j$. We denote the set of all measurement orderings for f in q as $\Omega_{f,M,q}$.

The fault signatures and measurement orderings can be automatically computed from a temporal causal graph representation that is derived from the system model, using a forward propagation algorithm to predict qualitative effects of faults on measurements and their possible sequences of deviations [Mosterman and Biswas, 1999; Daigle, 2008].

Given a sequence of observable events, i.e., measurement deviations and controlled mode changes, the fault isolation task consists of matching event sequences to hypothesized fault candidates. We define a candidate as a hypothesized fault and a hypothesized system mode.

Definition 3 (Candidate)**.** A *candidate* c is defined as $c = (f_i, q_i)$, where $f_i \in F$ is a hypothesized fault, and $q_i \in Q$ is a hypothesized current mode. The set of all candidates is denoted as C.

We wish to find candidates that are *consistent* with the sequence of observed events. A *diagnosis* is a collection of candidates that are consistent with the observations provided to the diagnoser after the time of fault occurrence, t_f.

Definition 4 (Diagnosis)**.** At time $t \geq t_f$, a *diagnosis* $d \subseteq C$ is a set of candidates consistent with the observations made on the system during the interval $[t_f, t]$.

Fault isolation is performed incrementally, as new events are received. At each new event, the current diagnosis is reduced by eliminating candidates that are inconsistent with the new event, given the previous sequence of events. Ideally, the diagnosis will eventually reduce to a unique candidate.

3 Event-based Fault Modeling

In order to characterize diagnosability in our framework, we first need to define what it means for a candidate to be consistent with a sequence of observable events. We do this by modeling the possible sequences of measurement deviations that

faults may cause in different modes as event traces. Candidate traces are then formed by a special composition of these individual traces to account for the interleavings of events caused by mode changes.

For a specific fault and mode, the combination of all fault signatures and relative measurement orderings yields all the possible ways a fault can manifest. We denote each of these possibilities as a *fault trace*.

Definition 5 (Fault Trace)**.** A *fault trace* for a fault f over measurements M in mode q, denoted by $\lambda_{f,M,q}$, is a string of length $\leq |M|$ that includes, for every $m \in M$ that will deviate due to f in q, a fault signature $\sigma_{f,m,q}$, such that the sequence of fault signatures satisfies $\Omega_{f,M,q}$.

Note that the definition implies that fault traces are of maximal length, i.e., a fault trace includes deviations for all measurements affected by the fault. We group the set of all fault traces into a *fault language*. The *fault model*, defined by a *finite automaton*, concisely represents the fault language.

Definition 6 (Fault Language)**.** The *fault language* of a fault $f \in F$ with measurement set M in mode q, denoted by $L_{f,M,q}$, is the set of all fault traces for f over measurements M in q.

Definition 7 (Fault Model)**.** The *fault model* for a fault $f \in F$ with measurement set M in mode q, is the finite automaton that accepts exactly the language $L_{f,M,q}$, and is given by $\mathcal{L}_{f,M,q} = (S, s_0, \Sigma, \delta, A)$ where S is a set of states, $s_0 \in S$ is an initial state, Σ is a set of events, $\delta : S \times \Sigma \rightarrow S$ is a transition function, and $A \subseteq S$ is a set of accepting states.

The finite automata representation allows for the composition of the fault signatures and relative measurement orderings into fault models. The possible fault signatures and measurement orderings can be composed automatically to form the fault models based on the synchronization operation [Daigle *et al.*, 2007a].

We need to define the candidate language in order to formally characterize consistency of candidates. Unlike fault traces, traces for candidates must contain both controlled mode change events and measurement deviation events. We denote the set of possible measurement deviation events as Σ_M, and the set of mode change events as Σ_Q.

When a controlled mode change occurs during fault isolation, the system model is updated, and a new nominal reference for symbol generation is computed. When a new measurement deviates in the new mode, current hypothesized candidates must match the predictions for these candidates in the new mode, ignoring previously deviated measurements, in order to still be consistent. There may be different possible modes of fault occurrence, depending on the history of control actions, therefore, the set of consistent candidates depends also on the expected mode of fault occurrence. Given this, we can now define a *candidate trace*. In the following, we denote the mode transition function of the system by μ.

Definition 8 (Candidate Trace)**.** An event trace $\lambda = \sigma$ is a *candidate trace* for $c = (f_i, q_i)$ and initial mode of fault occurrence q_0, if $\sigma \sqsubseteq \lambda' \in L_{f_i,M,q_i}$ where $q_i = \mu(f_i, q_0)$. An event trace $\lambda = \lambda_i \sigma_{i+1}$ is a *candidate trace* for $c = (f_i, q_{i+1})$ and initial mode of fault occurrence q_0, if λ_i is a candidate trace for (f_i, q_i), and if $\sigma_{i+1} \in \Sigma_Q$ then $\mu(\sigma_{i+1}, q_i) = q_{i+1}$, or if $\sigma_{i+1} \in \Sigma_M$ then $q_i = q_{i+1}$ and $\sigma_{i+1} \sqsubseteq \lambda' \in L_{f_i,M-M_i,q_{i+1}}$. A candidate trace for c with initial mode q_0 is denoted as λ_{c,q_0}.

In other words, given a candidate trace, an extension of that trace by a measurement deviation event will also be a candidate trace for the same candidate, if the deviation is consistent with the candidate for the new mode (i.e., it is consistent with the fault language in the new mode). An extension of the trace by a mode change event, however, will be a candidate trace for a different candidate, namely, the one defined by changing the mode of the old candidate to the new mode.

Clearly, there may be an infinite number of candidate traces because controlled mode changes may keep occurring indefinitely. However, we are only concerned with *maximal* traces, i.e., those for which all measurements that will deviate in the current mode have deviated (as with fault traces).

Definition 9 (Maximal Candidate Trace)**.** A candidate trace λ_{c,q_0} for $c = (f_i, q_i)$ is *maximal* if $L_{f,M-M_i,q_i} = \varnothing$, where M_i is the set of deviated measurements for λ_{c,q_0}.

Now, we can define the language of a candidate c with respect to an initial mode of fault occurrence q_0, L_{c,M,q_0} as the set of maximal candidate traces for c starting in q_0.

Definition 10 (Candidate Language)**.** The *candidate language* for candidate c, measurements M, and initial mode of fault occurrence q_0, denoted as L_{c,M,q_0}, is the set of all maximal candidate traces λ_{c,q_0}.

The candidate language consists of all consistent maximal traces for the candidate. A maximal trace is consistent with a candidate if the mode of the candidate can be reached via the sequence of controlled mode changes in the trace, and the measurement deviations within the trace match the fault in the intermediate modes.

4 Diagnosability

Diagnosability is an important property of a system, because it enables us to make guarantees about the unique isolation of faults. We first provide definitions of *distinguishability* and *diagnosability* and then describe how these notions are captured in our event-based framework.

If two candidates will always produce different effects, we say they are distinguishable. For hybrid systems, we must define distinguishability with respect to an initial expected mode at the point of fault occurrence, as with candidate traces.

Definition 11 (Distinguishability)**.** For an expected mode $q \in Q$ at the point of fault occurrence, a candidate c_i is distinguishable from a candidate c_j, denoted by $c_i \nsim_q c_j$, if for any possible sequence of controlled mode changes, c_i always eventually produces effects on the measurements that c_j cannot.

Candidate languages essentially capture the effects produced on the measurements for candidates, and thus characterize consistency of candidates with observed effects. Therefore, candidate languages can be used to establish distinguishability within our framework.

Lemma 1. *For an expected mode $q_0 \in Q$ at the point of fault occurrence, a candidate c_i is* distinguishable *from a candidate c_j given measurements M and possible modes Q, if there does not exist a pair of candidate traces $\lambda_{c_i,q_0} \in L_{c_i,M,q_0}$ and $\lambda_{c_j,q_0} \in L_{c_j,M,q_0}$ such that $\lambda_{c_i} \sqsubseteq \lambda_{c_j}$.*

Proof. Assume c_i is not distinguishable from c_j, i.e., $c_i \sim_{q_0} c_j$ for initial mode of fault occurrence q_0. Then, by definition, starting in mode q_0, there must exist a maximal candidate trace by c_i that c_j can also produce. Therefore, there must exist some maximal candidate trace for c_i, i.e., some $\lambda_{c_i,q_0} \in L_{c_i,M,q_0}$, and some sequence of events for c_j that is not distinct from λ_{c_i,q_0}. So, λ_{c_i,q_0} must be a candidate trace λ_{c_j,q_0} for c_j. Therefore, if $c_i \sim_{q_0} c_j$ then there exits some $\lambda_{c_i,q_0} \in L_{c_i,M,q_0}$ and $\lambda_{c_j,q_0} \in L_{c_j,M,q_0}$ such that $\lambda_{c_i,q_0} \sqsubseteq \lambda_{c_j,q_0}$. By the contrapositive, if there does not exist $\lambda_{c_i,q_0} \in L_{c_i,M,q_0}$ and $\lambda_{c_j,q_0} \in L_{c_j,M,q_0}$ such that $\lambda_{c_i,q_0} \sqsubseteq \lambda_{c_j,q_0}$, then $c_i \nsim_{q_0} c_j$. □

Since candidate traces include mode change events, the candidate languages cover all possible sequences of controlled mode change events interleaved with measurement deviations. Therefore, checking distinguishability is equivalent to checking for common traces. So, if a maximal candidate trace, which is a sequence of controlled mode change events and measurement deviation events, for some candidate is a prefix for a second candidate, then if the first candidate occurs and produces that trace, the candidates cannot be distinguished, because no more measurements will deviate (since the trace is maximal).

In our framework, a *system* can be defined as follows.

Definition 12 (System)**.** A *system* $\mathcal{S}$ is defined as $(F, M, Q, L_{F,M,Q})$, where $F = \{f_1, f_2, \ldots, f_n\}$ is a set of faults, $M = \{m_1, m_2, \ldots, m_p\}$ is a set of measurements, $Q = \{q_1, q_2, \ldots, q_r\}$ is a set of modes, and $L_{F,M,Q}$ is the set of fault languages for each fault in each mode, i.e., $L_{F,M,Q} = \{L_{f,M,q} : f \in F, q \in Q\}$.

Using distinguishability, we obtain the following notion of diagnosability for a hybrid system.

Definition 13 (Diagnosability)**.** A system $\mathcal{S} = (F, M, Q, L_{F,M,Q})$ is *diagnosable* if for all c_i and c_j and possible modes of fault occurrence $q_0 \in Q$, where $|c_i| \leq l$ and $|c_j| \leq l$, $c_i \nsim_{q_0} c_j$.

If the system is diagnosable, then every two candidates are distinguishable using the measurements in M. So, each sequence of measurement deviations and controlled mode changes we observe can be eventually linked to a diagnosis with a unique candidate. Hence, we can uniquely isolate all candidates of interest. If the system is not diagnosable, then ambiguities may remain after fault isolation, i.e., after all possible measurement deviations have been observed.

The definition of diagnosability allows making guarantees about fault isolation. Although controlled mode change events affect the diagnosis, since the diagnoser has no control over which controlled mode change events are issued, we cannot, in general, make any restrictions about when a mode change event will be issued. Thus, diagnosability in this sense is conservative. It may be possible, however, to avoid ambiguous diagnosis results if certain mode changes are blocked or executed. We define this as Q-diagnosability.

Definition 14 (Q-diagnosability)**.** A system $\mathcal{S} = (F, M, Q, L_{F,M,Q})$ is *Q-diagnosable* if for all c_i and c_j and possible modes of fault occurrence $q_0 \in Q$, where $c_i \sim_{q_0} c_j$, then for every (maximal) λ_{c_i,q_0} where $\lambda_{c_i,q_0} \sqsubseteq \lambda_{c_j,q_0}$, either there exists some sequence of controlled mode changes λ_Q where $\lambda_{c_i,q_0}\lambda_Q$ is not maximal for any candidate, or for every $\lambda_{c_k}\lambda_Q = \lambda_{c_i,q_0}$ where λ_Q is a sequence of controlled mode changes, λ_{c_k} is not maximal for any candidate.

If the system is Q-diagnosable, then for any trace that violates diagnosability, there is some sequence of controlled mode changes that can be applied such that the new trace is no longer maximal, i.e., more measurement deviations will occur, or for every partial trace that can become the violating trace via a sequence of controlled mode changes, the partial trace is not maximal. The first case corresponds to executing controlled mode changes to ensure more measurement deviations will occur. The second case corresponds to blocking a sequence of controlled mode changes such that we never encounter the violating trace in the first place.

5 Diagnoser Design

We construct from our fault models an event-based diagnoser, which is an extended form of a finite automaton. If our system is diagnosable, we can construct a diagnoser that uniquely isolates all candidates. If not, the constructed diagnoser will give ambiguous results for some maximal traces. But, if the system is Q-diagnosable, the ambiguous results can be avoided. We wish to use the diagnoser to help determine system diagnosability. The goal of the event-based diagnoser is, given a sequence of measurement deviation events and controlled mode change events, to determine which faults are consistent with the observed sequence. We define formally a *diagnoser* in our framework.

Definition 15 (Diagnoser)**.** A *diagnoser* for a fault set F, measurements M, and modes Q, is defined as $\mathcal{D}_{F,M,Q} = (S, I, \Sigma, \delta, A, D, Y)$ where S is a set of states, $I \subseteq S$ is set of initial states, Σ is a set of events, $\delta : S \times \Sigma \to S$ is a transition function, $A \subseteq S$ is a set of accepting states, $D \subseteq 2^C$ is a set of diagnoses, and $Y : S \to D$ is a diagnosis map.

A diagnoser is a finite automaton extended by a set of diagnoses and a diagnosis map. The initial states correspond to possible starting modes at the point of fault occurrence. A diagnoser takes events as inputs, which correspond to measurement deviations $\sigma \in \Sigma_M$ and controlled mode changes $\sigma \in \Sigma_Q$. From the current state, a measurement deviation event causes a transition to a new state. The diagnosis for that new state represents the set of candidates that are consistent with the sequence of events seen up to the current point in time, i.e., it encodes the results that hypothesis generation and refinement would obtain.

The accepting states of the diagnoser correspond to a fault isolation result. We say that a diagnoser *isolates* a candidate if it accepts all possible valid traces for the candidate and the accepting states map to diagnoses containing the candidate.

Definition 16 (Isolation)**.** A diagnoser $\mathcal{D}_{F,M,Q}$ *isolates* a candidate c if it accepts all $\lambda \in L_{c,M,q_0}$ for all nominal $q_0 \in Q$, and for each $s \in A$ that accepts a $\lambda \in L_{c,M,q_0}$, $c \in Y(s)$.

The notion of isolation gives us an indication of correctness of our diagnosers. If our diagnoser isolates all candidates, then it covers all possible observable fault traces, and, therefore, is constructed correctly. We also would like to achieve unique isolation of candidates, which is a stronger notion of isolation. For unique isolation, we require that the diagnoser isolates candidate c, but also that the corresponding accepting states uniquely determine c. This means that the diagnoser will accept all valid maximal candidate traces, but also that each trace will uniquely identify a single candidate.

Definition 17 (Unique Isolation)**.** A diagnoser $\mathcal{D}_{F,M,Q}$ *uniquely isolates* a candidate c if it isolates c and for each $s \in A$ that accepts some $\lambda_c \in L_{c,M,q_0}$, $\{c\} = Y(s)$.

Unique isolation relates to diagnosability, so it can provide us with guarantees about the ambiguity of the diagnosis results. If we can design a diagnoser that isolates all candidates of interest, then by examining the diagnoser we can determine if it uniquely isolates all candidates, and if so, that the system is diagnosable. If not diagnosable, we can also use the diagnoser to determine which traces result in ambiguities, and if possible, avoid those traces by permitting or prohibiting certain controlled mode changes during isolation, i.e., achieve Q-diagnosability.

Ultimately, we would like to systematically construct a diagnoser for a hybrid system $\mathcal{S}$ that isolates all possible candidates. Further, we would like to show that if $\mathcal{S}$ is diagnosable, then this diagnoser *uniquely* isolates all candidates. To do this, we use individual diagnosers for each fault-mode pair, and provide a composition operator to simultaneously compose all the individual diagnosers to a global diagnoser that isolates all the valid candidates.

First, we construct a diagnoser, $\mathcal{D}^*_{\{f\},M,q}$ for each single fault f within each mode q from $\mathcal{L}_{f,M,q}$.

Definition 18 ($\mathcal{D}^*_{\{f\},M,q}$)**.** Given fault f and mode q for measurements M, with $\mathcal{L}_{f,M,q} = (S,\ s_0,\ \Sigma,\ \delta,\ A)$, $\mathcal{D}^*_{\{f\},M}$ is defined as $(S,\ s_0,\ \Sigma,\ \delta,\ A',\ \{\{(f,q)\}\},\ Y)$, where $Y(s) = \{(f,q)\}$ for all $s \in S$, and $A' = A$ if $S \neq \{s_0\}$, or $A' = \{s_0\}$ otherwise.

We simultaneously compose each of the individual diagnosers $\mathcal{D}_{\{f\},M,q}$. In incremental consistency checking, we project out measurements that have already deviated to obtain the set of consistent candidates for a new observation. For a diagnoser, the state-based form of the measurement projection operation on traces is formalized using *boundaries* and *boundary transition functions*.

Definition 19 (Boundary)**.** The *boundary* for a state s and deviated measurements M_i, $B_{M_i}(s)$, is defined as the set of all states $\delta(\lambda, s)$ such that λ contains only measurement deviation events corresponding to those in M_i.

The boundary for a state s is basically the set of states that may be transitioned to from s via a trace λ consisting of only events for measurements that have already deviated, i.e., measurements corresponding to the events for traces in the history of the state. Using the notion of a boundary, we define a *boundary transition function* with respect to a set of deviated measurements.

Definition 20 (Boundary Transition Function)**.** The *boundary transition function* for an event σ, state s, and set of deviated measurements M_i, denoted as $\delta_{M_i}(\sigma, s)$, is a transition function that maps σ and s to some state s', such that $s' = \varnothing$ if the cardinality of $\{\delta(\sigma, s_B) : s_B \in B_{M_i}(s)\}$ is not 1, or s' is the single element in $\{\delta(\sigma, s_B) : s_B \in B_{M_i}(s)\}$, otherwise.

In other words, $\delta_{M_i}(\sigma, s)$ returns the unique state that can be reached from a boundary state of s via σ, or $\varnothing$ if there are no states that can be reached or if the state is not unique. Because of the way the $\mathcal{D}^*_{\{f\},M,q}$ diagnosers are computed, the reachable state will always be unique or null, because traces with the same set of measurements map to the same state. In the following, we denote the measurements that have deviated in a state s as $M(s)$.

We now describe a composition operator, Π, that simultaneously combines the $\mathcal{D}^*_{\{f\},M,q}$ for each possible (f,q) pair. We split the mode set Q into nominal modes Q_N and faulty modes Q_F.

Definition 21 (Π Composition)**.** Given the set of all k (f,q) diagnosers, $\mathbb{D} = \{\mathcal{D}^*_{\{f\},M,q} : f \in F, q \in Q\}$, $\mathcal{D}^*_{F,M,Q} \triangleq \Pi(\mathbb{D})$, where

- $I = \{(s_{0,1}, s_{0,2}, \ldots, s_{0,k}, q, (\varnothing, q)) : q \in Q_N\}$
- $\Sigma = \Sigma_1 \cup \Sigma_2 \cup \ldots \cup \Sigma_k \cup \Sigma_Q$
- $\delta(\sigma, (s_{i,1}, s_{i,2}, \ldots, s_{i,k}, q_i, d_i)) = (s_{i+1,1}, s_{i+1,2}, \ldots, s_{i+1,k}, q_{i+1}, d_{i+1})$, where $\sigma \in \Sigma_Q$, all $s_{i+1,j} = s_{i+1,j}$, $q_{i+1} = \mu(\sigma, q_i)$, and $d_{i+1} = \{(f, \mu(\sigma,q)) : \mu(\sigma,q) \neq \varnothing \wedge (f,q) \in d_i\}$
- $\delta(\sigma, (s_{i,1}, s_{i,2}, \ldots, s_{i,k}, q_i, d_i)) = (s_{i+1,1}, s_{i+1,2}, \ldots, s_{i+1,k}, q_{i+1}, d_{i+1})$, where $\sigma \in \Sigma_M$, $q_{i+1} = q_i$, $M_i = M((s_{i,1}, s_{i,2}, \ldots, s_{i,k}, q_i, d_i))$, $s_{i+1,j} = s_{i,j}$ if $\delta_{M_i,j}(\sigma, s_{i,j}) = \varnothing$, or $\delta_{M_i,j}(\sigma, s_{i,j})$ otherwise, and $d_{i+1} = \{(f,q) \in d_i : \sigma \sqsubseteq \lambda \in L_{f,M-M_i,q}\} \neq \varnothing$
- S is the set of all s reachable through δ from some $s_0 \in I$
- A is the set of all $s_i = (s_{i,1}, s_{i,2}, \ldots, s_{i,k}, q_i, d_i) \in S$ where there exists some $s_{i,j} \in s_i$, with some $s_{B,j} \in B_{M(s_i)}(s_{i,j})$ where $s_{B,j} \in A_j$, such that $Y_j(s_{B,j}) \subseteq\in Y(s_i)$
- D is the set of all d_i in each $(s_{i,1}, s_{i,2}, \ldots, s_{i,k}, q_i, d_i) \in S$
- $Y((s_{i,1}, s_{i,2}, \ldots, s_{i,k}, q_i, d_i)) = d_i$

Theorem 1. *The diagnoser $\mathcal{D}^*_{F,M,Q}$ isolates all valid candidates.*

Proof. Assume initial mode of fault occurrence q_0, candidate c, and trace $\lambda = \sigma_1\sigma_2\ldots\sigma_k \in L_{c,M,q_0}$. By the definition of a candidate trace, σ_1 is a candidate trace for $c' = (f_i, \mu(f, q_0))$ if $\sigma_1 \sqsubseteq \lambda' \in L_{f,M,\mu(f,q_0)}$. Therefore,

$(f, \mu(f, q_0)) \in h_{F,M_i}(\sigma_1)$, so by definition of $\wedge_l$, the resultant diagnosis will contain $(f, \mu(f, q_0))$, so by definition of δ, the corresponding state is in S. Assume λ_i is a candidate trace for $c' = (f_i, q_i)$ and has a corresponding state $s \in S$. Then if $\sigma_{i+1} \in \Sigma_Q$, $\lambda_i \sigma_{i+1}$ is a candidate trace for $(f_i, \mu(\sigma_{i+1}, q_i))$ and by definition of δ has a corresponding state $s \in S$ and the associated diagnosis has $(f_i, \mu(\sigma_{i+1}, q_i))$. If $\sigma_{i+1} \notin \Sigma_Q$, then $\lambda_i \sigma_{i+1}$ is a candidate trace for (f_i, q_i) if $\sigma_{i+1} \sqsubseteq \lambda' \in L_{f,M,\mu(f,q_0)}$ and therefore by definition of a hypothesis set, $(f_i, q_i) \in h_{F,M_i}(\sigma_{i+1})$, so by definition of $\wedge_l$, the diagnosis will contain (f_i, q_i) and by definition of δ, will have a corresponding state in S. Therefore, there is a state for any valid candidate trace. Given a state $s \in S$ with a trace that is maximal for $c = (f_i, q_i)$, the substate of s that corresponds to a state in $\mathcal{D}^*_{f,M,q_i}$ must have no measurement deviations possible from its boundary, otherwise the trace would not be maximal, and thus the boundary must contain an accepting state. □

Further, we can show that if the system $\mathcal{S}$ is diagnosable, then the diagnoser uniquely isolates all candidates.

Theorem 2. *A system $\mathcal{S} = (F, M, Q, L_{F,M,Q})$ is diagnosable if and only if $\mathcal{D}^*_{F,M,Q}$ uniquely isolates all valid candidates.*

Proof. Assume $\mathcal{S}$ is diagnosable. Assume a c and $\lambda \in L_{c,M,Q}$. $\mathcal{D}^*_{F,M,Q}$ isolates c, so must have corresponding accepting state s with $c \in Y(s)$. Since $\mathcal{S}$ is diagnosable, there cannot be a $c'$where c and c' are not distinguishable, by definition of diagnosability. So, there cannot be some common subtrace λ that maps to an accepting state that has both c' and c. So, $\mathcal{D}^*_{F,M,Q}$ uniquely isolates all c. Assume $\mathcal{D}^*_{F,M,Q}$ uniquely isolates all c. Then each possible fault trace λ has an accepting state s where $c \in Y(s)$. Thus, there cannot be some c', with trace λ' that reaches the same state, otherwise c' is in $Y(s)$. Therefore, c and c' are distinguishable, so $\mathcal{S}$ is diagnosable. Thus $\mathcal{S}$ is diagnosable if and only if $\mathcal{D}^*_{F,M,Q}$ uniquely isolates all c. □

6 Case Study

We apply the diagnosability framework to the Advanced Diagnostics and Prognostics Testbed (ADAPT) deployed at NASA Ames [Poll *et al.*, 2007]. The testbed is functionally representative of a spacecraft's electrical power system, and consists of three subsystems: (*i*) power generation, which includes two battery chargers, (*ii*) power storage, which consists of three sets of lead-acid batteries, and (*iii*) power distribution, which consists of a number of relays and circuit breakers, two inverters, and various DC and AC loads.

We consider a subset of ADAPT to demonstrate our approach, which includes a lead-acid battery, two relays, and two DC loads. The battery is modeled by an electric circuit equivalent described in [Daigle, 2008] (see Fig. 2). The battery supplies voltage to the relays through a parallel connection, which in turn supply power to the two DC loads. The selected measurements are the battery voltage, $V_B(t)$, and the currents through the relays, $I_{L1}(t)$ and $I_{L2}(t)$, i.e., $M = \{I_{L1}, I_{L2}, V_B\}$.

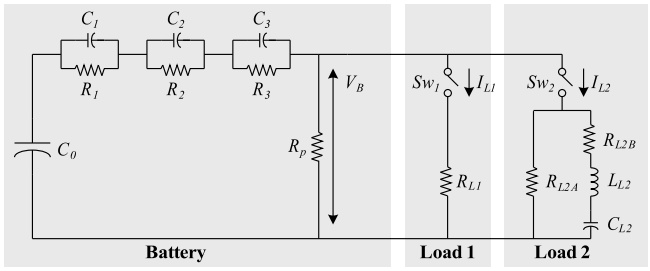


Figure 2: Electrical circuit equivalent for the selected subsystem.

We consider faults in the battery, loads, relays, and sensors. Common battery faults include loss of charge and resistance increases brought about by battery use and age, which manifest as a side effect of the chemical reactions. Loss of charge is represented by a capacitance decrease, C_0^-, and an increase in internal losses by R_1^+. Faults can occur in the system loads, and these are represented by increases or decreases in their resistance values, R_{L1} and R_{L2A}. For the sensors, we consider bias faults, which produce abrupt changes in the measured values manifesting as constant offsets. Sensor faults are labeled by the measured quantity they represent, e.g., V_B^+ represents a bias fault in the battery voltage sensor. We represent discrete faults in Sw_1 and Sw_2 by fault events α and β, respectively, where a subscript of 0 indicates a stuck-off fault, and a subscript of 1 indicates a stuck-on fault.

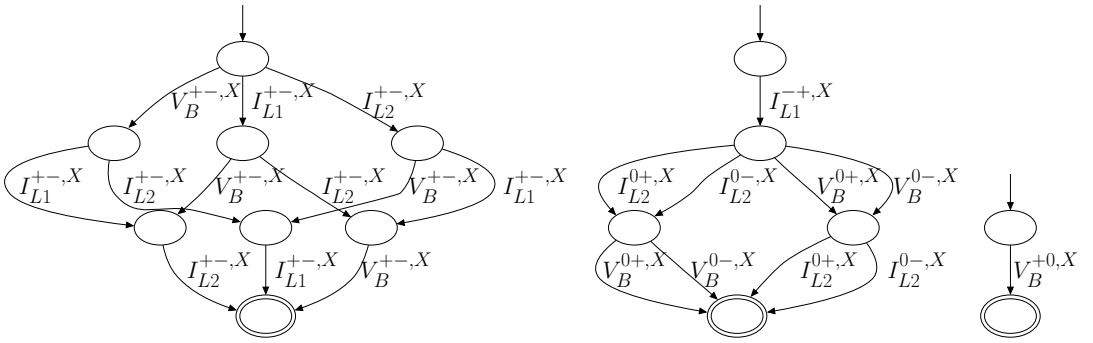
6.1 Diagnosability Analysis

We denote the system mode as q_{ij} and a controlled mode change to q_{ij} as $\sigma_{q_{ij}}$, where i is the mode of Sw_1, and j is the mode of Sw_2. We allow controlled mode changes that switch the system from any one controlled mode to another, i.e., $\Sigma_Q = \{\sigma_{q_{00}}, \sigma_{q_{01}}, \ldots, \sigma_{q_{11}}\}$. We restrict discrete faults to only occurring from expected modes where a deviation would be produced, e.g., α_1 would not produce any deviations if it occurred in a mode where Sw_1 was already on.

The fault signatures and relative measurement orderings for the chosen faults are given in Table 1 for selected modes (q_{**} indicates the signatures and orderings are valid for any mode). The nonlinearities in the battery introduce ambiguity in the qualitative signatures, and this is denoted by the $\star$ symbol, e.g., a signature of $0\star$ may manifest as $0+$ or $0-$. Since the sensors are not part of any feedback loops in the system, sensor faults affect only the measurement provided by the sensor. The other measurements are not affected, and so the corresponding fault signatures are denoted by 00, indicating no change in the measurement from expected behavior.

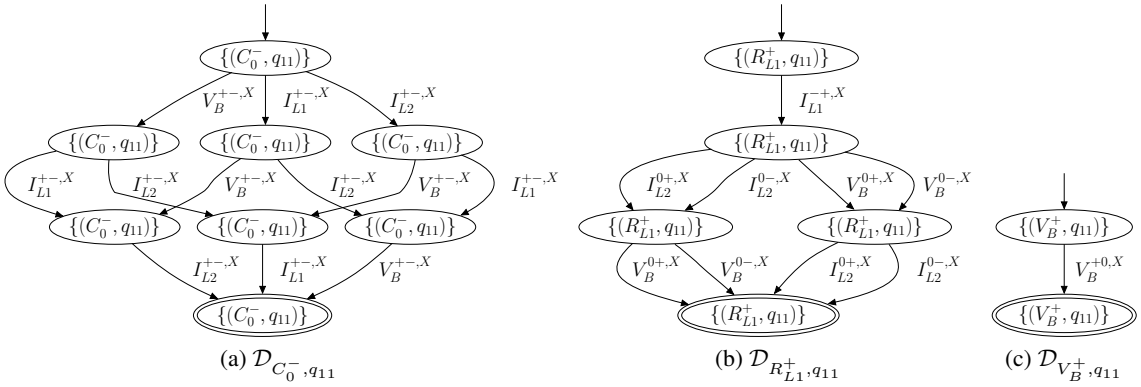
Selected fault models for ADAPT are shown in Fig. 3. Consider the fault model $\mathcal{L}_{R_{L1}+,q_{11}}$, shown in Fig. 3b. Note that the M subscript is dropped in the notation. From the orderings, the current through Load 1 must be the first to deviate, followed by the Load 2 current and battery voltage in any order. The direction of the changes in $I_{L2}(t)$ and $V_B(t)$ are unknown so both possibilities are represented. The individual diagnosers for the same faults are shown in Fig. 4.

Given any one mode, the system is diagnosable. After at



(a) $\mathcal{L}_{C_0^-,q_{11}}$ (b) $\mathcal{L}_{R_{L1}^+,q_{11}}$ (c) $\mathcal{L}_{V_B^+,q_{11}}$

Figure 3: Selected fault models for ADAPT.



(a) $\mathcal{D}_{C_0^-,q_{11}}$ (b) $\mathcal{D}_{R_{L1}^+,q_{11}}$ (c) $\mathcal{D}_{V_B^+,q_{11}}$

Figure 4: Selected individual diagnosers for ADAPT.

Fault	V_B	I_{L1}	I_{L2}	Measurement Orderings
(V_B^+, q_{**})	+0,X	00,X	00,X	$V_B \prec I_{L1}, V_B \prec I_{L2}$
(V_B^-, q_{**})	−0,X	00,X	00,X	$V_B \prec I_{L1}, V_B \prec I_{L2}$
(I_{L1}^+, q_{**})	00,X	+0,X	00,X	$I_{L1} \prec V_B, I_{L1} \prec I_{L2}$
(I_{L1}^-, q_{**})	00,X	−0,X	00,X	$I_{L1} \prec V_B, I_{L1} \prec I_{L2}$
(I_{L2}^+, q_{**})	00,X	00,X	+0,X	$I_{L1} \prec V_B, I_{L2} \prec I_{L1}$
(I_{L2}^-, q_{**})	00,X	00,X	−0,X	$I_{L1} \prec V_B, I_{L2} \prec I_{L1}$
(C_0^-, q_{11})	+−,X	+−,X	+−,X	$\varnothing$
(R_1^+, q_{11})	0−,X	0−,X	0−,X	$\varnothing$
(R_{L1}^+, q_{11})	0⋆,X	−+,X	0⋆,X	$I_{L1} \prec V_B, I_{L1} \prec I_{L2}$
(R_{L1}^-, q_{11})	0⋆,X	+−,X	0⋆,X	$I_{L1} \prec V_B, I_{L1} \prec I_{L2}$
(R_{L2A}^+, q_{11})	0⋆,X	0⋆,X	−+,X	$I_{L2} \prec V_B, I_{L2} \prec I_{L1}$
(R_{L2A}^-, q_{11})	0⋆,X	0⋆,X	+−,X	$I_{L2} \prec V_B, I_{L2} \prec I_{L1}$
$(\alpha_0, q_{\alpha_0 1})$	0⋆,X	−⋆,Z	0⋆,X	$I_{L1} \prec V_B, I_{L1} \prec I_{L2}$
$(\alpha_1, q_{\alpha_1 1})$	0⋆,X	+⋆,N	0⋆,X	$I_{L1} \prec V_B, I_{L1} \prec I_{L2}$
$(\beta_0, q_{1\beta_0})$	0⋆,X	0⋆,X	−⋆,Z	$I_{L2} \prec V_B, I_{L2} \prec I_{L1}$
$(\beta_1, q_{1\beta_1})$	0⋆,X	0⋆,X	+⋆,N	$I_{L2} \prec V_B, I_{L2} \prec I_{L1}$

Table 1: Fault Signatures and Relative Measurement Orderings for the ADAPT Subsystem

most two measurement deviations, a unique candidate can be isolated. However, over all modes, the system is not diagnosable. Fig. 5 gives a partial diagnoser for the system that illustrates this property, with $F = \{C_0^-, R_{L1}^+\}$ and initial mode q_{11} with $\sigma_{q_{01}}$ and $\sigma_{q_{11}}$ being the only controlled mode change events. If $I_{L1}^{+-,X}\sigma_{q_{01}}$ occurs, we reach an accepting state that corresponds to a diagnosis with multiple candidates. After that event, both C_0^- and R_{L1}^+ are consistent. Since the state is accepting, it is possible that no new measurement deviations will occur to distinguish the faults. The resistance fault will have no visible effects on the rest of the measurements in this mode, because the source of the deviations is cut off, so we would have to wait infinitely long to verify R_{L1}^+ as the true fault. Therefore, the system is not diagnosable. We can see, however, that the system is Q-diagnosable. If we prevent $\sigma_{q_{01}}$ from occurring, or change back to q_{11} if it does occur, more measurements will deviate and we can distinguish the candidate uniquely. Additional diagnosability results that include multiple faults and autonomous mode changes, as well as diagnosis experiments, are reported in [Daigle, 2008].

7 Conclusions

We presented a systematic framework to create event-based diagnosers for hybrid systems. Using the diagnosers, diagnosability of the system can be analyzed. We introduced the notion of Q-diagnosability, in which unique isolation can be achieved if certain controlled mode changes are prevented or

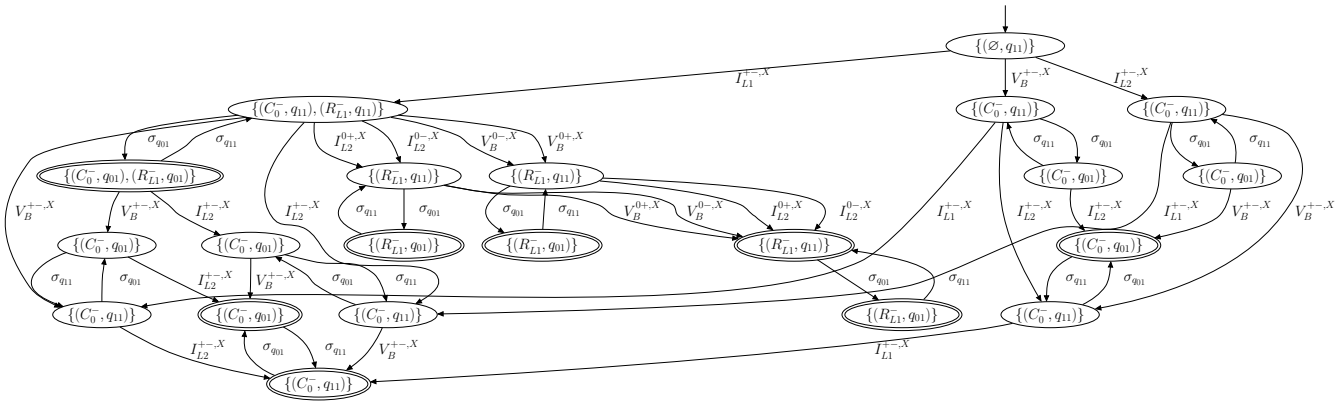


Figure 5: Partial hybrid diagnoser for $F = \{C_0^-, R_{L1}^+\}$ and initial mode q_{11}.

executed during the fault isolation stage. We applied the technique to analyze the diagnosability of a subset of the ADAPT system.

Acknowledgments

This work was supported in part by grants NSF CNS-0615214, NASA USRA 08020-013, NASA NRA NNX07AD12A, and NSF CNS-0347440.

References

[Bayoudh *et al.*, 2006] M. Bayoudh, L. Traveé-Massuyès, and X. Olive. Hybrid systems diagnosability by abstracting faulty continuous dynamics. In *Proc. of the 17th International Workshop on Principles of Diagnosis*, pages 9–15, 2006.

[Benedetto *et al.*, 2007] Di Benedetto, S. Di Gennaro, and A. D'Innocenzo. Diagnosability verification for hybrid automata and durational graphs. In *Proc. of the 46th IEEE Conference on Decision and Control*, pages 1789–1794, December 2007.

[Biswas *et al.*, 2003] G. Biswas, G. Simon, N. Mahadevan, S. Narasimhan, J. Ramirez, and G. Karsai. A robust method for hybrid diagnosis of complex systems. In *Proc. of the 5th Symposium on Fault Detection, Supervision and Safety for Technical Processes*, pages 1125–1131, June 2003.

[Cordier *et al.*, 2006] M.-O. Cordier, L. Travé-Massuyès, and X. Pucel. Comparing diagnosability in continuous and discrete-event systems. In *Proc. of the 17th International Workshop on Principles of Diagnosis.*, pages 55–60, 2006.

[Daigle *et al.*, 2007a] M. Daigle, X. Koutsoukos, and G. Biswas. Fault diagnosis of continuous systems using discrete-event methods. In *Proc. of the 46th IEEE Conference on Decision and Control*, pages 2626–2632, December 2007.

[Daigle *et al.*, 2007b] M. J. Daigle, X. D. Koutsoukos, and G. Biswas. Distributed diagnosis in formations of mobile robots. *IEEE Trans. on Robotics*, 23(2):353–369, April 2007.

[Daigle *et al.*, 2008] M. Daigle, X. Koutsoukos, and G. Biswas. An integrated approach to parametric and discrete fault diagnosis in hybrid systems. In *HSCC 2008*, volume 4981 of *LNCS*, pages 614–617. Springer-Verlag, 2008.

[Daigle, 2008] M. Daigle. *A Qualitative Event-based Approach to Fault Diagnosis of Hybrid Systems*. PhD thesis, Vanderbilt University, 2008.

[Meseguer *et al.*, 2008] J. Meseguer, V. Puig, and T. Escobet. Fault diagnosis using a timed discrete event approach based on interval observers. In *Proc. of the 17th IFAC World Congress*, 2008.

[Mosterman and Biswas, 1999] P.J. Mosterman and G. Biswas. Diagnosis of continuous valued systems in transient operating regions. *IEEE Trans. on Systems, Man and Cybernetics, Part A*, 29(6):554–565, 1999.

[Narasimhan and Biswas, 2007] S. Narasimhan and G. Biswas. Model-based diagnosis of hybrid systems. *IEEE Trans. on Systems, Man and Cybernetics, Part A*, 37(3):348–361, May 2007.

[Poll *et al.*, 2007] S. Poll *et al*. Evaluation, selection, and application of model-based diagnosis tools and approaches. In *AIAA Infotech@Aerospace 2007 Conference and Exhibit*, May 2007.

[Sampath *et al.*, 1995] M. Sampath, R. Sengupta, S. Lafortune, K. Sinnamohideen, and D.C. Teneketzis. Diagnosability of discrete-event systems. *IEEE Trans. on Automatic Control*, 40(9):1555–1575, September 1995.

[Travé-Massuyès *et al.*, 2006] L. Travé-Massuyès, T. Escobet, and X. Olive. Diagnosability analysis based on component-supported analytical redundancy relations. *IEEE Trans. on Systems, Man and Cybernetics, Part A*, 36(6):1146–1160, November 2006.

[Zad *et al.*, 2003] S. Hashtrudi Zad, R.H. Kwong, and W.M. Wonham. Fault diagnosis in discrete-event systems: framework and model reduction. *IEEE Trans. on Automatic Control*, 48(7):1199–1212, July 2003.