Qualitative Fault Isolation of Hybrid Systems: A Structural Model Decomposition-Based Approach

Anibal Bregon[1], Matthew Daigle[2], and Indranil Roychoudhury[3]

[1] *Department of Computer Science, University of Valladolid, Valladolid, Spain*
anibal@infor.uva.es

[2] *NASA Ames Research Center, Moffett Field, California, 94035, USA*
matthew.j.daigle@nasa.gov

[3] *Stinger Ghaffarian Technologies Inc., NASA Ames Research Center, Moffett Field, California, 94035, USA*
indranil.roychoudhury@nasa.gov

ABSTRACT

Quick and robust fault diagnosis is critical to ensuring safe operation of complex engineering systems. A large number of techniques are available to provide fault diagnosis in systems with continuous dynamics. However, many systems in aerospace and industrial environments are best represented as hybrid systems that consist of discrete behavioral modes, each with its own continuous dynamics. These hybrid dynamics make the on-line fault diagnosis task computationally more complex due to the large number of possible system modes and the existence of autonomous mode transitions. This paper presents a qualitative fault isolation framework for hybrid systems based on structural model decomposition. The fault isolation is performed by analyzing the qualitative information of the residual deviations. However, in hybrid systems this process becomes complex due to possible existence of observation delays, which can cause observed deviations to be inconsistent with the expected deviations for the current mode in the system. The great advantage of structural model decomposition is that (*i*) it allows to design residuals that respond to only a subset of the faults, and (*ii*) every time a mode change occurs, only a subset of the residuals will need to be reconfigured, thus reducing the complexity of the reasoning process for isolation purposes. To demonstrate and test the validity of our approach, we use an electric circuit simulation as the case study.

1. INTRODUCTION

The development of robust and efficient fault diagnosis techniques plays an important role in complex engineering systems. A large number of fault diagnosis techniques have been developed during the last few decades for continuous systems. However, nowadays, many engineering systems are modeled as *hybrid systems* that have both continuous and discrete-event dynamics. For such systems, the large number of possible system modes with different dynamics and the existence of autonomous mode transitions significantly increases the complexity of the fault diagnosis problem.

Different proposals exist in the literature for hybrid systems diagnosis, focusing on either hybrid modeling, such as hybrid automata (Henzinger, 2000; Rienmüller, Bayoudh, Hofbaur, & Travé-Massuyès, 2009; Bayoudh, Travé-Massuyès, & Olive, 2008), hybrid state estimation (Hofbaur & Williams, 2004), or a combination of on-line state tracking and residual evaluation (Benazera & Travé-Massuyès, 2009). However, in all those approaches, the proposed solutions involve modeling and pre-enumeration of the set of *all* possible system-level discrete modes, which grows exponentially with the number of switching components. Both steps are computationally very expensive or even infeasible for hybrid systems with a large number of complex interacting subsystems.

One of the solutions to avoid the mode pre-enumeration problem consists of building hybrid system models in a *compositional* way, where discrete modes are defined at a local level (e.g., at the component level), and the system-level mode is defined implicitly by the local component-level modes. This allows the modeler to focus on the discrete behavior only at the component level, and the pre-enumeration of all the system-level modes can be avoided (Narasimhan & Brownston, 2007; Trave-Massuyes & Pons, 1997). Additionally, building models in a compositional way facilitates reusability and maintenance, and allows the validation of the com-

ponents individually before they are composed to create the system-level hybrid model.

In previous work (Daigle, Bregon, & Roychoudhury, 2015), we proposed a compositional modeling approach for hybrid systems, where models consist of user-defined components. A component is constructed by defining a set of discrete modes, with a different set of mathematical constraints describing the continuous dynamics in each mode. Within this framework, we defined algorithms for efficient causality assignment and re-assignment upon mode changes. For a given system mode, structural model decomposition (Roychoudhury, Daigle, Bregon, & Pulido, 2013) is used to construct minimal submodels for residual generation, and, based on efficient causality reassignment, can be efficiently reconfigured upon mode changes.

We demonstrated in (Daigle, Bregon, & Roychoudhury, 2015) how the minimal submodels can be used for efficient residual generation over the different modes of the system. In this paper, we extend this framework with a qualitative approach for online fault isolation of hybrid systems. Our approach assumes only single faults occur in the system and we choose to deal with deviations in system parameters (i.e. parametric faults) only and not discrete faults. The approach works by abstracting qualitatively the transients of residual deviations and comparing those with the predicted fault transients. Unlike previous approaches based on this methodology (P. J. Mosterman & Biswas, 1999; Daigle, Koutsoukos, & Biswas, 2009; Narasimhan & Biswas, 2007), we make use of structural model decomposition to decrease the complexity of the diagnosis task. In hybrid systems, mode changes typically modify the predicted fault transients, and, further, observation delays (e.g., due to delays in signal filtering within a fault detection algorithm, or communication delays) can cause that the observed transient may be consistent with a previous mode of the system. As it has been discussed by other authors (Narasimhan & Biswas, 2007), both issues significantly complicate the reasoning process. Using structural model decomposition, the associated complexity can be reduced greatly because (i) it allows the design of residuals that respond to only a subset of the faults (Bregon et al., 2014); and (ii) every time a mode change occurs, only a subset of the residuals will need to be reconfigured (Daigle, Bregon, & Roychoudhury, 2015). Using an electrical circuit as a case study, we demonstrate that our approach can correctly isolate faults in hybrid systems even if the system transitions among different mode changes and presents observation delays during the isolation process.

The paper is organized as follows. Section 2 summarizes the compositional modeling approach and introduces the case study. Section 3 presents the problem we solve in this paper. Section 4 describes the qualitative fault isolation approach for hybrid systems. Section 5 demonstrates the approach for the electrical case study. Section 6 reviews the related work and current approaches for hybrid systems fault diagnosis and puts our work into context. Finally, Section 7 concludes the paper.

2. COMPOSITIONAL HYBRID SYSTEMS MODELING

As we have mentioned, in (Daigle, Bregon, & Roychoudhury, 2015) we proposed a compositional approach for hybrid systems modeling, in which system models are made up of a set of user-defined components, where each component is defined by a set of discrete modes, with a different set of constraints describing the continuous dynamics of the component in each mode. In this section, we summarize the main details of the hybrid system modeling framework and structural model decomposition approach. For additional details, we refer the reader to (Daigle, Bregon, & Roychoudhury, 2015).

2.1. System Modeling

At the basic level, the continuous dynamics of a component in each mode are modeled using a set of *variables* and a set of *constraints*. A constraint is defined as follows:

Definition 1 (Constraint)**.** A constraint c is a tuple (ε_c, V_c), where ε_c is an equation involving variables V_c.

A component is defined by a set of constraints over a set of variables. The constraints are partitioned into different sets, one for each component mode. A component is then defined as follows:

Definition 2 (Component)**.** A *component* κ with n discrete modes is a tuple $\kappa = (V_\kappa, \mathcal{C}_\kappa)$, where V_κ is a set of variables and $\mathcal{C}_\kappa$ is a set of constraints sets, where $\mathcal{C}_\kappa$ is defined as $\mathcal{C}_\kappa = \{C_\kappa^1, C_\kappa^2, \ldots, C_\kappa^n\}$, with a constraint set, C_κ^m, defined for each mode $m = \{1, \ldots, n\}$.

To illustrate our proposal, throughout the paper we will use a circuit example, shown in Fig. 1. The components of the circuit are a voltage source, V, two capacitors, C_1 and C_2, two inductors, L_1 and L_2, two resistors, R_1 and R_2, and two switches, Sw_1 and Sw_2, as well as components for series and parallel connections. Sensors measure the current or voltage in different locations (i_3, v_8, and i_{11}, as indicated in Fig. 1). Because each switch has two modes (on and off), there are four total modes in the system. The components of the circuit are defined in Table 1.

Example 1. Consider the component Sw_2 (κ_{10}). It has two modes: *on* (represented as mode 2 in Table 1) and *off* (represented as mode 1). In the *off* mode, it has three constraints setting each of its currents (i_9, i_{10}, i_{11}) to 0. In the *on* mode, it has also three constraints, setting the three currents equal to each other and establishing that the voltages sum up (it acts like a series connection when in the on mode).

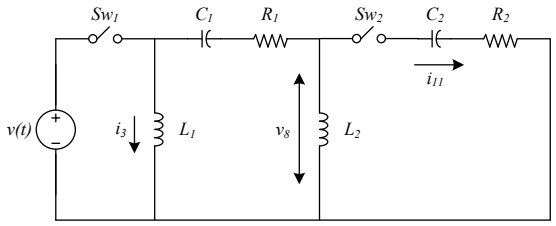


Figure 1. Electrical circuit running example.

Table 1. Components of the electrical circuit.

Component	Mode	Constraints
κ_1: V	1	$v_1{=}u_v$
κ_2: Sw_1	1	$i_1{=}0$
		$i_2{=}0$
	2	$i_1{=}i_2$
		$v_1{=}v_2$
κ_3: Parallel Connection$_1$	1	$v_2{=}v_3$
		$v_2{=}v_4$
		$i_2{=}i_3 + i_4$
κ_4: L_1	1	$\dot{f}_3{=}v_3$
		$i_3{=}f_3/L_1$
		$f_3{=}\int_{t_0}^{t} \dot{f}_3$
κ_5: Series Connection$_1$	1	$i_4{=}i_5$
		$i_4{=}i_6$
		$i_4{=}i_7$
		$v_4{=}v_5 + v_6 + v_7$
κ_6: R_1	1	$v_5{=}i_5 * R_1$
κ_7: C_1	1	$\dot{q}_6{=}i_6$
		$v_6{=}q_6/C_1$
		$q_6{=}\int_{t_0}^{t} \dot{q}_6$
κ_8: Parallel Connection$_2$	1	$v_7{=}v_8$
		$v_7{=}v_9$
		$i_7{=}i_8 + i_9$
κ_9: L_2	1	$\dot{f}_8{=}v_8$
		$i_8{=}f_8/L_2$
		$f_8{=}\int_{t_0}^{t} \dot{f}_8$
κ_{10}: Sw_2	1	$i_9{=}0$
		$i_{10}{=}0$
		$i_{11}{=}0$
	2	$i_9{=}i_{10}$
		$i_9{=}i_{11}$
		$v_9{=}v_{10} + v_{11}$
κ_{11}: R_2	1	$v_{10}{=}i_{10} * R_2$
κ_{12}: C_2	1	$\dot{q}_{11}{=}i_{11}$
		$v_{11}{=}q_{11}/C_2$
		$q_{11}{=}\int_{t_0}^{t} \dot{q}_{11}$
κ_{13}: Current Sensor$_{11}$	1	$i_{11}^*{=}i_{11}$
κ_{14}: Voltage Sensor$_8$	1	$v_8^*{=}v_8$
κ_{15}: Current Sensor$_3$	1	$i_3^*{=}i_3$

We define a system model as a set of components:

Definition 3 (Model)**.** A *model* $\mathcal{M} = \{\kappa_1, \kappa_2, \ldots, \kappa_k\}$ is a finite set of k components for $k \in \mathbb{N}$.

Example 2. The model of the electrical system is made up of the components detailed in Table 1, i.e., $\mathcal{M} = \{\kappa_1, \kappa_2, \ldots, \kappa_{15}\}$. For each component, the variables and constraints are defined for each component mode.

The set of variables for a model, $V_\mathcal{M}$, is the union of all the component variable sets, i.e., for d components, $V_\mathcal{M} = V_{\kappa_1} \cup V_{\kappa_2} \cup \ldots \cup V_{\kappa_d}$. $V_\mathcal{M}$ consists of five disjoint sets, namely, the set of state variables, $X_\mathcal{M}$; the set of parameters, $\Theta_\mathcal{M}$; the set of inputs (variables not computed by any constraint), $U_\mathcal{M}$; the set of outputs (variables not used to compute any other variables), $Y_\mathcal{M}$; and the set of auxiliary variables, $A_\mathcal{M}$. Parameters, $\Theta_\mathcal{M}$, include explicit model parameters that are used in the model constraints (e.g., fault parameters). Auxiliary variables, $A_\mathcal{M}$, are additional variables that are used to simplify the structure of the equations.

Example 3. In the circuit model, we have $X_\mathcal{M} = \{f_3, q_6, f_8, q_{11}\}$, $\Theta_\mathcal{M} = \{L_1, R_1, C_1, L_2, R_2, C_2\}$, $U_\mathcal{M} = \{u_v\}$, and $Y_\mathcal{M} = \{i_3^*, i_{11}^*, v_8^*\}$. Remaining variables belong to $A_\mathcal{M}$. Here, the * superscript is used to denote a measured value of a physical variable, e.g., $i_3 \in X_\mathcal{M}$ is the current and $i_3^* \in Y_\mathcal{M}$ is the measured current.

The interconnection structure of the model is captured using shared variables between components, i.e., components κ_i and κ_j are connected if $V_{\kappa_i} \cap V_{\kappa_j} \neq \varnothing$.

Example 4. In the circuit model, component κ_5 (Series Connection$_1$) is connected to κ_3 (Parallel Connection$_1$) through i_4, to κ_6 (R_1) through i_5 and v_5, to κ_7 (C_1) through i_6 and v_6, and κ_8 (Parallel Connection$_2$) through i_7 and v_7.

In our work, a fault is the cause of an unexpected, persistent deviation of the system behavior from the acceptable nominal behavior. To simplify our approach, we link faults only to the set of parameters $\Theta_\mathcal{M}$, i.e., no discrete faults are considered. More formally, a fault is defined as follows.

Definition 4 (Fault)**.** A fault, denoted as *f*, is a persistent deviation of exactly one parameter $\theta \in \Theta_\mathcal{M}$ of the system model $\mathcal{M}$ from its nominal value.

The model constraints, $C_\mathcal{M}$, are a union of the component constraints over all modes, i.e., $C_\mathcal{M} = \mathcal{C}_{\kappa_1} \cup \mathcal{C}_{\kappa_2} \cup \ldots \cup \mathcal{C}_{\kappa_d}$. Constraints are exclusive to components, that is, a constraint $c \in C_\mathcal{M}$ belongs to exactly one $\mathcal{C}_\kappa$ for $\kappa \in \mathcal{M}$.

To refer to a particular mode of a model we use the concept of a *mode vector*. A mode vector $\mathbf{m}$ specifies the current mode of each of the components of a model. So, the constraints for a mode $\mathbf{m}$ are denoted as $C_\mathcal{M}^\mathbf{m}$.

Example 5. Consider a model with five components, then if $\mathbf{m} = [1, 1, 3, 2, 1]$, it indicates that components κ_1, κ_2, and κ_5 use constraints of their mode 1, component κ_3 use

constraints of its mode 3, and component κ_4 use constraints of its mode 2.

For shorthand, we will refer to the modes only of the components with multiple modes. So, for the circuit, we will refer only to components κ_2 and κ_{10}, and we will have four possible mode vectors, $[1\ 1], [1\ 2], [2\ 1]$, and $[2\ 2]$.

The switching behavior of each component can be defined using a finite state machine or a similar type of control specification. For the purposes of this paper, we view the switching behavior as a black box where the mode change event is given, and refer the reader to many of the approaches already proposed in the literature for modeling the switching behavior (Henzinger, 2000; P. Mosterman & Biswas, 2000).

2.2. Structural Model Decomposition for Hybrid Systems

In our framework, we use structural model decomposition to generate submodels for the purpose of computing residuals, i.e., the difference between observed and predicted system behavior, which are then used for diagnosis. The main advantage that structural model decomposition provides, in contrast to using a global model for residual generation, is that each residual is designed to respond to only a subset of the faults, thus decreasing the complexity of diagnosis. Further, it allows the diagnosis task to be distributed, improving scalability (Bregon et al., 2014). We will show later, in Section 4, the specific advantages that structural model decomposition provides for diagnosis of hybrid systems.

In order to derive submodels, we need to assign causality to the system. Given a constraint c, belonging to a specific mode of a specific component, the notion of a *causal assignment* is used to specify a possible computational direction, or *causality*, for the constraint c. This is done by defining which $v \in V_c$ is the dependent variable in equation ε_c. For a given mode, we have the set of causal assignments over the entire model in that mode, and with that we can compute the minimal submodels, using the `GenerateSubmodel` algorithm described in our previous work (Roychoudhury et al., 2013). The algorithm finds a submodel, which computes a set of local outputs given a set of local inputs, by searching over the causal model. It starts at the local inputs, and propagates backwards through the causal constraints, finding which constraints and variables must be included in the submodel. When possible, causal constraints are inverted in order to take advantage of local inputs. Additional information and the pseudocode are provided in (Roychoudhury et al., 2013).

In the context of residual generation, we set the local output set to a single measured value, and the local inputs to all other measured values and the (known) system inputs. That is, we exploit the analytical redundancy provided by the sensors in order to find minimal submodels to estimate values of sensor outputs. In this framework, we consider one submodel per sensor, each producing estimated values for that sensor. Assuming that the set of sensors does not change from mode to mode, we will always have one submodel per sensor. Since the set of constraints changes from mode to mode, the set of submodels will change as well, however, by taking advantage of causality information, reconfiguring the submodels can be done efficiently (Daigle, Bregon, & Roychoudhury, 2015).

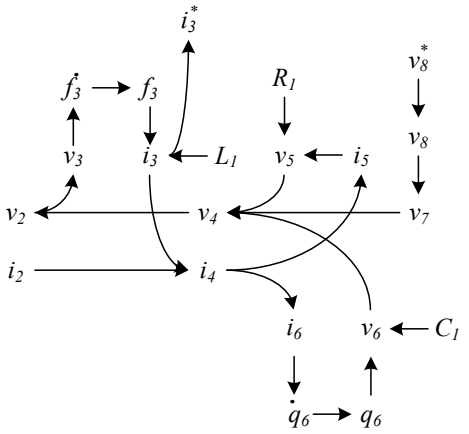
Example 6. Submodels can be represented visually using a graph notation, where vertices correspond to variables, and edges correspond to constraints with causal assignments, i.e., a directed edge from v_i to v_j means that v_j is computed using v_i. The submodel graphs for the circuit in mode $\mathbf{m} = [1\ 2]$ are shown in Fig. 2, and in mode $\mathbf{m} = [2\ 1]$ in Fig. 3. For example, consider i_{11}^*. In the first mode, it is computed using the measurements v_8^* and i_3^* as inputs. The variable i_2 is effectively an input; it is set to 0 since Sw_1 is off. Here, only a fault in L_2 will propagate to i_{11}^*. In the second mode, Sw_2 is off, and so i_{11} is set to 0, and the submodel contains only i_{11} and i_{11}^*, and these variables are decoupled from all faults.

3. PROBLEM FORMULATION

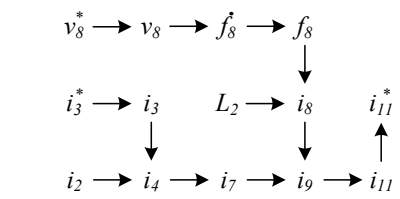
Our qualitative fault diagnosis approach (Daigle, Roychoudhury, & Bregon, 2015) works by reasoning over observations of how measurements deviate from expected nominal behavior. These observations are formed from a qualitative abstraction of residual signal deviations. Residuals are computed as the difference between predicted nominal, $\hat{y}(t)$, and measured, $y(t)$, system variables, i.e. a residual $r(t)$ is computed as $r(t) = y(t) - \hat{y}(t)$. Predicted system variables $\hat{y}(t)$ are computed using the minimal submodels as described in the previous section. Fault detection works by determining statistically significant nonzero deviations in the residuals. Residual deviations are then abstracted into a symbolic representation to form *fault signatures*. These symbols are computed from the residuals using symbol generation, as described in (Daigle, Roychoudhury, & Bregon, 2015). Finally, the predicted signatures are compared with observed signatures in order to isolate faults.

In the context of hybrid systems, the structure of the residual generators changes from mode to mode, causing the set of fault signatures to also change. Observing mode change events can help to match the observations to both the fault and the mode in which they occurred. Further, if there is a delay in the observation of fault signatures, then the mode in which the deviation actually occurred may not be the current mode in the system in which it was observed, and consequently, the fault signature for the estimated fault could mismatch the fault signature for the current mode. A hybrid system diagnosis algorithm must handle each of these challenges.

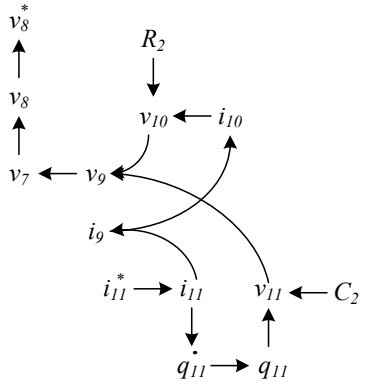
We restrict the problem to single faults.



(a) i_3^* submodel graph.

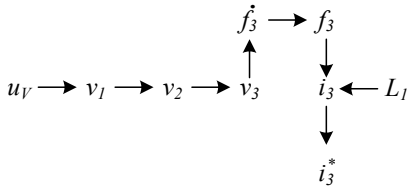


(b) i_{11}^* submodel graph.



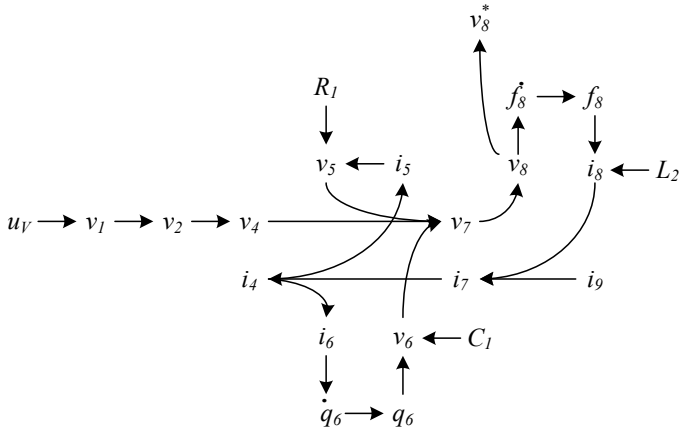
(c) v_8^* submodel graph.

Figure 2. Submodel graphs for $\mathbf{m} = [1\ 2]$.



(a) i_3^* submodel graph.

(b) i_{11}^* submodel graph.



(c) v_8^* submodel graph.

Figure 3. Submodel graphs for $\mathbf{m} = [2\ 1]$.

Assumption 1. Only single faults occur in the system.

Thus, we define a diagnosis as follows.

Definition 5 (Diagnosis)**.** For a system with fault set F, a *diagnosis* is a fault $f \in F$ that is consistent with a given finite sequence of observations. A set of diagnoses is denoted as D.

In our diagnosis definition we appeal the Principle of Parsimony as stated by (Reiter, 1987), meaning that a diagnosis is a conjecture that some minimal set of components are faulty.

The diagnosis problem can then be formally defined as follows.

Problem 1. For a system with fault set F, given a finite sequence of observations O, find the set of diagnoses $D \subseteq F$ that is consistent with O.

4. QUALITATIVE FAULT ISOLATION FOR HYBRID SYSTEMS

Generally speaking, for the purposes of diagnosis, we consider an observation to be an event observed at a particular time.

Definition 6 (Observation)**.** An *observation* is a tuple (e, t), where e is an observed event and t is the time of observation.

Table 2. Fault Signatures for global model of the electrical system.

Mode	$\mathbf{m} = [1\ 2]$				$\mathbf{m} = [2\ 1]$			
Fault	$r_{i_{11}^*}$	$r_{i_3^*}$	$r_{v_8^*}$	Orderings	$r_{i_{11}^*}$	$r_{i_3^*}$	$r_{v_8^*}$	Orderings
C_1^-	0−	0+	0−	$\varnothing$	00	00	−+	$r_{v_8^*} \prec r_{i_3^*}, r_{v_8^*} \prec r_{i_{11}^*}$
C_2^-	0+	0−	−+	$r_{v_8^*} \prec r_{i_3^*}, r_{v_8^*} \prec r_{i_{11}^*}$	00	00	00	$\varnothing$
L_1^-	−+	+−	−⋆	$\varnothing$	00	+0	00	$r_{i_3^*} \prec r_{i_{11}^*}, r_{i_3^*} \prec r_{v_8^*}$
L_2^-	−+	0−	−⋆	$r_{i_{11}^*} \prec r_{i_3^*}, vr_{v_8^*} \prec r_{i_3^*}$	00	00	−⋆	$r_{v_8^*} \prec r_{i_3^*}, r_{v_8^*} \prec r_{i_{11}^*}$
R_1^+	0+	0−	0+	$\varnothing$	00	00	+−	$r_{v_8^*} \prec r_{i_3^*}, r_{v_8^*} \prec r_{i_{11}^*}$
R_2^+	0+	0−	−+	$r_{v_8^*} \prec r_{i_3^*}, r_{v_8^*} \prec r_{i_{11}^*}$	00	00	00	$\varnothing$

We consider two types of events: (*i*) fault signature events and (*ii*) mode change events. Section 4.1 reviews the event-based fault modeling framework based on the concepts of fault signatures (Daigle et al., 2009), and extends it to hybrid systems. Following that, Section 4.2 describes how diagnostic reasoning can be performed under this new framework in the presence of mode changes.

4.1. Event-based Fault Modeling

The basis of the qualitative fault isolation approach is the concept of a fault signature.

Definition 7 (Fault Signature)**.** A *fault signature* for a fault f and residual r in mode m, denoted by $\sigma_{f,r,m}$ is a pair of symbols $s_1 s_2$ representing potential qualitative changes in magnitude and slope of r caused by f at the point of the occurrence of f in mode m. The set of all fault signatures for a fault f over residuals R in mode m is denoted as $\Sigma_{f,R,m}$.

When a fault occurs, it produces a transient in the observed behavior with respect to the predicted nominal behavior, observed as changes in the residual signal (P. J. Mosterman & Biswas, 1999). These changes are formulated as qualitative changes (+, −, or 0) in residual magnitude and slope.

Changes are observed in each residual that is a function of the fault. Thus, when a fault occurs, we observe a *sequence* of fault signatures. *Relative residual orderings* define a partial order of signatures for a given fault, and thus define all the possible fault signature sequences that can be produced by a fault.

Definition 8 (Relative Residual Ordering)**.** A *relative residual ordering* for a fault f and residuals r_i and r_j in mode m is a tuple (r_i, r_j), denoted by $r_i \prec_{f,m} r_j$, representing that f always manifests in r_i before r_j in mode m. The set of all orderings for a fault f over residuals R in mode m is denoted as $\Omega_{f,R,m}$.

Example 7. Table 2 shows the fault signatures for two modes of the circuit system for the global model residuals. For example, in mode $\mathbf{m} = [1\ 2]$, C_2^- will cause a −+ in $r_{v_8^*}$, i.e., a decrease in magnitude and increase in slope. On $r_{i_3^*}$ it will cause 0−, i.e, no change in magnitude and an increase in slope. In $\mathbf{m} = [2\ 1]$, however, C_2 is disconnected from the circuit and so cannot affect any of the residuals.

Table 3. Fault Signatures for minimal submodels of the electrical system.

Mode	$\mathbf{m} = [1\ 2]$			$\mathbf{m} = [2\ 1]$		
Fault	$r_{i_{11}^*}$	$r_{i_3^*}$	$r_{v_8^*}$	$r_{i_{11}^*}$	$r_{i_3^*}$	$r_{v_8^*}$
C_1^-	00	0+	00	00	00	−+
C_2^-	00	00	−0	00	00	00
L_1^-	00	+−	00	00	+0	00
L_2^-	−0	00	00	00	00	−⋆
R_1^+	00	0−	00	00	00	+−
R_2^+	00	00	+0	00	00	00

Example 8. Table 3 shows the fault signatures for the circuit for the same two modes for the local submodel residuals. Since residuals are computed independently, no orderings can be declared. Consider again the fault C_2^-. In $\mathbf{m} = [1\ 2]$, it now affects only the residual for v_8^*, as it appears only in that local submodel (see Fig. 2). In fact, this is true for all faults - each is found in exactly one local submodel and so will affect exactly one residual, in either mode.

A single sequence of fault signatures is termed a *fault trace*.

Definition 9 (Fault Trace)**.** A *fault trace* for a fault f over a set of residuals R in mode m, denoted by $\lambda_{f,R,m}$, is a sequence of fault signatures that can be observed given the occurrence of f in mode m.

Fault traces are grouped into *fault languages*.[1]

Definition 10 (Fault Language)**.** The *fault language* for a fault f and residual set R in mode m, denoted by $L_{f,R,m}$, is the set of all fault traces for f over R in m.

[1] Fault languages can be automatically derived for certain classes of system models (Daigle, 2008), obtained via simulation, or obtained experimentally. In this work, we assume that the fault languages are given as input.

For the purposes of this paper, we assume that signatures and orderings are correctly observed.[2]

Assumption 2 (Correct Observation)**.** If a fault f occurs in mode m, then the observed fault trace will belong to $L_{f,R,m}$.

4.2. Hybrid Systems Diagnosis

For hybrid systems, fault signatures, residual orderings, fault traces, and fault languages are a function of the system mode. If the mode does not change between the point of fault occurrence and the diagnosis of the fault, then the problem reduces to the continuous systems case. Otherwise, we will observe some new trace that may not belong to any mode-specific fault language, i.e., it may be a trace that is composed of partial traces for a fault from the different modes encountered during diagnosis.

Example 9. For example, consider the global model residuals (Table 2). Assume that the system starts in $\mathbf{m} = [1\ 2]$ and R_1^+ occurs. Then we could observe $r^{0+}_{i^*_{11}}$, followed by $r^{0-}_{i^*_3}$. So far, this partial trace can be found as a prefix to a trace in $L_{R_1^+,R,[1\ 2]}$. Now, assume that the system moves to mode $\mathbf{m} = [2\ 1]$, now we would observe $r^{+-}_{v^*_8}$. This trace is not found in any mode-specific fault language.

Thus, the first challenge is that now observed fault traces may contain some subtraces corresponding to one mode, and other subtraces corresponding to other modes. Thus, the fault isolation reasoning must span over several potential mode changes. If we know the system mode, then we know which fault language corresponds to the predicted observations for each fault. If there are unobservable mode changes, this adds another layer of complexity, because we must not only diagnose which fault has occurred but also what mode the system is currently in. For the purposes of this paper, we make the following assumption.

Assumption 3 (Mode Change Observability)**.** All mode change events are observable.

Given Assumption 3, we can define mode change events as follows.

Definition 11 (Mode Change Event)**.** An event e_m represents the system changing from its current mode to mode m.

However, even if we know the current mode of the system, there is another layer of complexity to consider: *observation delay*. Specifically, in our framework, this corresponds to the observations of fault signatures being delayed. The difficulty is that the system may be in one mode, but when the observation arrives we have moved to a different system mode, and thus we do not know in which mode the observation was actually made.

[2]Relaxation of this assumption has been explored for continuous systems in (Daigle, Roychoudhury, & Bregon, 2014).

Algorithm 1 $D_{i+1} \leftarrow$ `FaultIsolation`$(D_i, \lambda_i, \sigma_{i+1}, M_\Delta)$

1: $D_{i+1} \leftarrow \varnothing$
2: **for all** $q \in M_\Delta$ **do**
3: **for all** $f \in D_i \cap F_{r,q}$ **do**
4: **if** $\sigma_{i+1} \in \Sigma_{f,r_{\sigma_{i+1}},m}$ **and** $\neg\exists r' \in (R - R_{\lambda_i})$ s.t. $r' \prec r_{\sigma_{i+1}} \in \Omega_{f,R-R_{\lambda_i},m}$ **then**
5: $D_{i+1} \leftarrow \{f\}$

Example 10. Consider again the previous example, in which the global model residuals are used, the system starts in $\mathbf{m} = [1\ 2]$ and R_1^+ occurs. Again, we observe $r^{0-}_{i^*_{11}}$, followed by $r^{0+}_{i^*_3}$, and then change to $\mathbf{m} = [2\ 1]$. Say that $r^{0-}_{v^*_8}$ occurred in the previous mode, but we only see get the observation now. This observation is not consistent with R_1^+ in $\mathbf{m} = [2\ 1]$.

Observation delay can manifest in different ways. For example, fault detection is usually performed by checking whether a residual crosses some threshold. To make this approach robust to noise, usually we check that the mean of the residual, computed over some small time window, has crossed that threshold. This means that the signal could actually cross the threshold in one mode, but the mean of the signal could cross only in the next mode. Thus, the observation of this signature is delayed. In practice, we can assume that observation delay is finite and bounded.

Assumption 4 (Bounded Observation Delay)**.** The delay of any observation is no greater than Δ.

Given our assumptions, the algorithm for a single step of fault isolation for hybrid systems is shown as Algorithm 1.[3] As inputs, it takes the current diagnosis, D_i, the previous sequence of fault signatures, λ_i, the new fault signature, σ_{i+1}, and the set of recent modes that falls within $[t - \Delta, t]$, M_Δ. The change from the continuous systems case is that we need to check signatures and orderings for each of the recent modes. If it is consistent in any of the modes, it must be added to D_{i+1}. Here, for a given mode m, we need to check only the subset of faults that are included in the current diagnosis and can actually affect this residual in this mode, denoted as $F_{r,m}$. To check consistency, we check that the predicted signature for the residual associated with σ_{i+1}, denoted as $r_{\sigma_{i+1}}$, can be found in the signature set for that fault and residual, and that the orderings, with respect to residuals that have not yet deviated (those in $R - R_{\lambda_i}$, where R_{λ_i} denotes the residuals associated with the trace λ_i), are not violated.

Algorithm 1 executes a single reasoning step, given a newly observed fault signature. This would be placed within a progressive monitoring algorithm, that keeps track of the current

[3]Because fault languages can become prohibitively large, we implement the fault isolation step directly using the signatures and orderings, which is more efficient (Daigle et al., 2009).

diagnosis, and computes the set of recent modes based on the times events are observed.

The complexity of the fault isolation algorithm is dependent on the number of faults, $|F|$, the number of residuals, $|R|$, and the number of modes, $|M|$. For the global model case, each time a new residual deviates, we must check, for each mode in M_Δ, whether each fault is consistent. The local submodel approach improves over the global model approach by simultaneously reducing both the effective $|R|$ and the effective $|M|$. The effective $|R|$ is decreased, because with structural model decomposition each fault affects only a subset of the residuals, so for each new residual deviation only a a subset of faults needs to be checked for consistency. The effective $|M|$ is reduced, because with structural model decomposition each residual reconfigures only based on a few local component modes, whereas for the global model each residual is dependent on the system-level modes (which increases exponentially with the number of switching components). Due to these properties of structural model decomposition, the approach scales at a significantly smaller rate as the system size increases than with the global model approach.

Example 11. Consider the residual $r_{i_3^*}$. For the global model residuals (Table 2), all 6 faults can affect this residual in $\mathbf{m} = [1\ 2]$, but in $\mathbf{m} = [2\ 1]$, only 1 fault affects it. So, if we are unsure of the mode in which the observation was actually made, all 6 faults must be considered in the conflict set. For the local submodel residuals (Table 3), 3 faults affect the residual in $\mathbf{m} = [1\ 2]$ and 1 in $\mathbf{m} = [2\ 1]$. In a system with more modes, this number increases at a much smaller rate than for the global model, due to the effects of the decomposition.

5. DEMONSTRATION OF APPROACH

In this section, we demonstrate the approach through some example scenarios using the circuit system. We consider two modes: one where Sw_1 is on and Sw_2 is off (i.e., $\mathbf{m} = [2\ 1]$), and one where Sw_1 is off and Sw_2 is on (i.e., $\mathbf{m} = [1\ 2]$). In all cases, the system starts in mode $\mathbf{m} = [2\ 1]$, switches to $\mathbf{m} = [1\ 2]$ at $t = 10$ s, and switches back to $\mathbf{m} = [2\ 1]$ at a later time, depending on the scenario. The complete fault candidate set considered for diagnosis is $\{C_1^-, R_1^+, L_1^-, C_2^-, R_2^+, L_2^-\}$. In each case, we compare the performance of the global model approach and the local submodel approach.

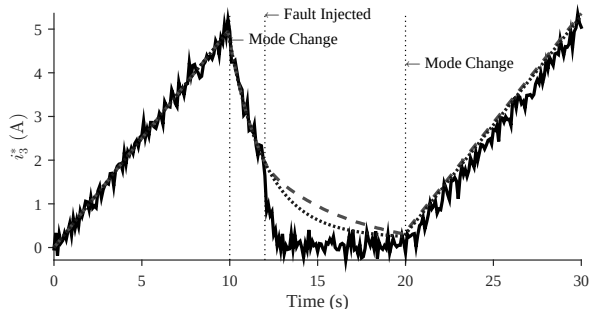
The symbol generation approach described in (Daigle et al., 2010) is used, which uses the Z-test for statistical fault detection and symbol generation. A window of samples is used to compute the mean, and thus can produce a delay that increases with window size. For the particular fault detector settings, we consider the bounded observation delay to be $\Delta = 5$ s.

Example 12 (R_1^+ fault)**.** In this scenario, an increase in R_1 is injected at $t = 12$ s. The measured and estimated values are shown in Fig. 4, which show that the behavior can be tracked through the mode changes during nominal operation. The residuals are shown in Fig. 5. In the global model residuals, we first observe $r_{v_8^*}^{0+}$ at $t = 12.2$ s (Fig. 5c), which can be due only to R_1^+ (see Table 2). We then observe at 12.3 s, $r_{i_{11}^*}^{0+}$ (Fig. 5b) and $r_{i_3^*}^{0-}$ (Fig. 5a), and the diagnosis remains the same. Since a mode change occurred within 5 s prior these observations, we must consider that the fault may have occurred in the previous mode ($\mathbf{m} = [2\ 1]$). However, none of these signatures are consistent with any fault in that mode, and so the diagnosis remains the same. In the local submodel residuals, we first observe $r_{i_3^*}^{0-}$ at $t = 12.3$ s (Fig. 5a), which, in this mode, is consistent only with R_1^+. In the previous mode, it is not consistent with any fault, and thus this is our diagnosis. At $t = 20.0$ s, a second mode change occurs, and, in this mode, $r_{v_8^*}$ will now respond to both these faults, and so $r_{v_8^*}^{+-}$ is observed (Fig. 5c), confirming the previous diagnosis.

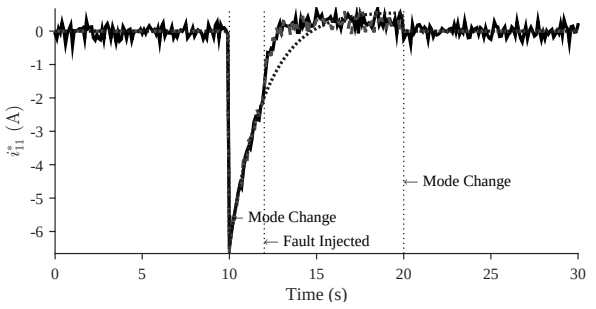
Example 13 (C_1^- fault)**.** In this scenario, a decrease in C_1 is injected at $t = 12$ s. The residuals are shown in Fig. 6. For the global model residuals, we first observe $r_{v_8^*}^{0-}$ at 12.5 s, which can be due to only C_1^- (Table 2). Next, a mode change occurs at 12.7 s. At 13.0 s, we observe $r_{i_3^*}^{0+}$, yet in this new mode we do not expect the fault to have any effect on $r_{i_3^*}$, i.e., this is a delayed observation from the previous mode, consistent still with C_1^-. No further residuals deviate. For the local submodel residuals, we observe first $r_{v_8^*}^{-+}$ at 12.7 s, which is when the mode changes after the fault injection. This signature is consistent with C_1^- occurring in this mode and L_2^- in this mode (Table 3). At 12.9 s, we observe $r_{i_3^*}^{0+}$, which is not expected in this mode, in fact it is delayed from the previous mode and consistent only with C_1^-, ruling out L_2^-, so C_1^- is the only diagnosis. No other residuals deviate, and so no more reasoning is performed.

6. RELATED WORK

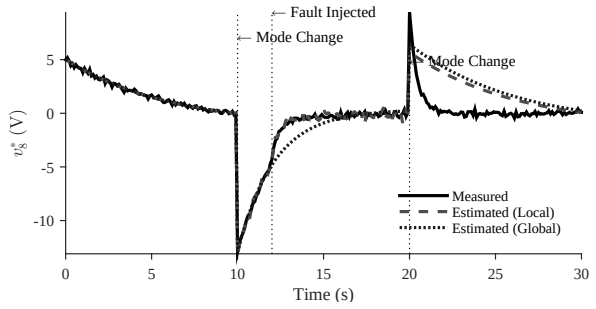
During the last decade or so, modeling and diagnosis for hybrid systems have been an important topic of researchers from both the FDI and DX communities. In the FDI community, several hybrid system diagnosis approaches have been developed. In (Cocquempot, El Mezyani, & Staroswiecki, 2004), parameterized ARRs are used. However, the approach is not suitable for systems with high nonlinearities or a large set of modes. In the DX community, some approaches have used different kind of automata to model the complete set of modes and transitions between them. In those cases, the main research topic has been hybrid system state estimation, which has been done using probabilistic (e.g., some kind of filter (Koutsoukos, Kurien, & Zhao, 2003) or hy-



(a) i_3^*.

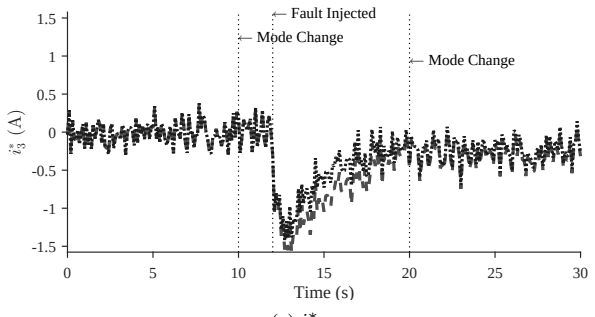


(b) i_{11}^*

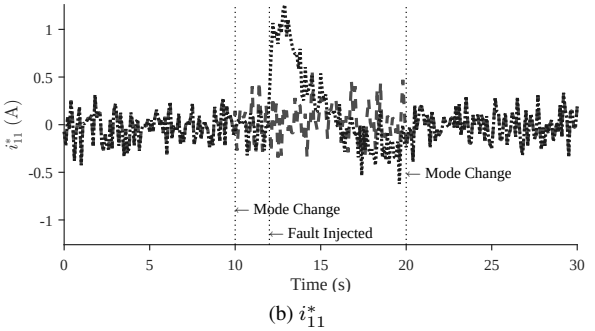


(c) v_8^*

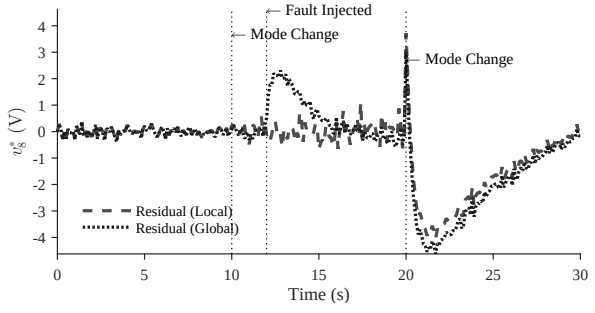
Figure 4. Measured and estimated values with an increase in R_1 at $t = 12$ s.



(a) i_3^*.



(b) i_{11}^*



(c) v_8^*

Figure 5. Residual values with an increase in R_1 at $t = 12$ s.

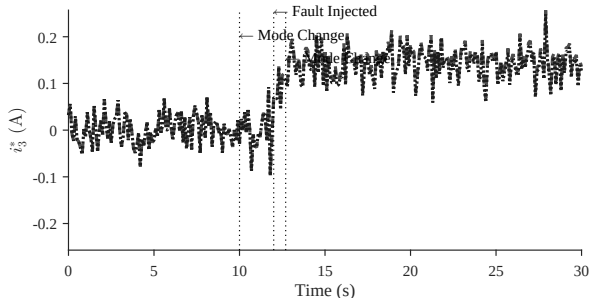
brid automata (Hofbaur & Williams, 2004)) or set-theoric approaches (Benazera & Travé-Massuyès, 2009).

Another solution has been to use an automaton to track the system mode, and then use a different technique to diagnose the continuous behavior (for example, using a set of ARRs for each mode (Bayoudh et al., 2008), or parameterized ARRs for the complete set of modes (Bayoudh, Travé-Massuyès, & Olive, 2009)). Nevertheless, one of the main difficulties regarding state estimation using these techniques is the need to pre-enumerate the set of possible system-level modes and mode transitions, which is difficult for complex systems. We avoid this problem by using a compositional approach.

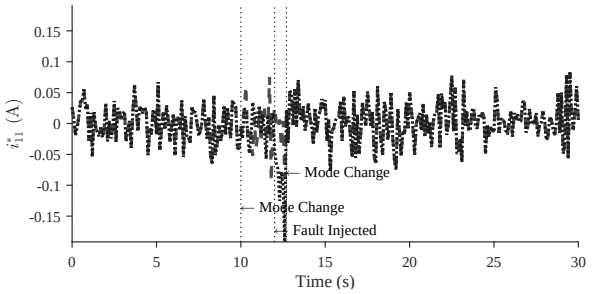
In (Alonso, Bregon, Alonso-González, & Pulido, 2013), the authors present a qualitative fault isolation approach for hybrid systems that is based on structural model decomposition. This approach, however, unlike ours, does not take into account observation delays. Moreover, the approach presented in (Alonso et al., 2013) is applicable only to systems that are modeled using hybrid bond graphs.

The focus of the research published in (Gaudel, Chanthery, & Ribot, 2015) is the treatment of both knowledge-based and observation-based uncertainty in health monitoring of hybrid systems. The diagnosis approach can reason with unobservable discrete events (e.g., faults), as well as false observations. However, unlike our generic formulation of hybrid systems, this work is restricted to systems modeled using the Hybrid Particle Petri Nets (HPPN) formalism.

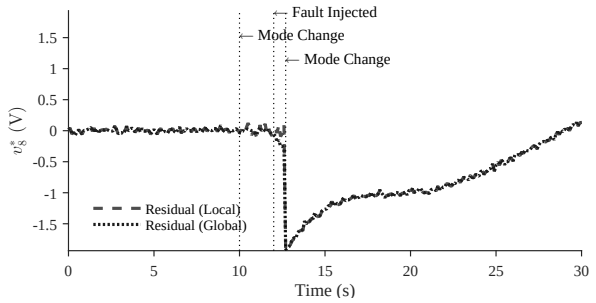
Finally, in (Bregon, Narasimhan, Roychoudhury, Daigle, & Pulido, 2013), the authors had developed an efficient model-based methodology for diagnosis that integrated structural model decomposition within the Hybrid Diagnosis Engine (HyDE) architecture (Narasimhan & Brownston, 2007). The



(a) i_3^*.



(b) i_{11}^*



(c) v_8^*

Figure 6. Residual values with a decrease in C_1 at $t = 12$ s.

HyDE architecture offers flexibility to choose the modeling paradigm and reasoning algorithms for diagnosis of hybrid systems. The authors show how the integration of the structural model decomposition reduces the computational complexity associated with the fault diagnosis of hybrid systems. In our paper, similar reduction in computational complexity of fault diagnosis is observed, further bolstering the support for using structural model decomposition for hybrid systems diagnosis.

7. CONCLUSIONS

In this work, we have developed a qualitative fault isolation approach for hybrid systems using structural model decomposition. Fault isolation is performed by analyzing the qualitative information of the residual signals. It has been proven that structural model decomposition can be used for hybrid systems fault isolation in the presence of observation delays, while the complexity of the isolation process can be reduced compared to a global model approach. The approach was demonstrated with a circuit system. In future work, we will further develop the hybrid systems diagnosis approach for discrete faults and for multiple fault diagnosis, and we will apply our approach to more complex systems. We will also show mathematically the improvement in the computational cost of the local algorithms. Finally, the assumption about unobservable mode changes occurring in the system can also be dropped, using the ideas developed in (Narasimhan & Biswas, 2007).

ACKNOWLEDGMENTS

This work has been funded by the Spanish MINECO DPI2013-45414-R grant and the NASA SMART-NAS project in the Airspace Operations and Safety Program of the Aeronautics Mission Directorate.

REFERENCES

Alonso, N. M., Bregon, A., Alonso-González, C. J., & Pulido, B. (2013). A common framework for fault diagnosis of parametric and discrete faults using possible conflicts. In *Advances in artificial intelligence* (pp. 239–249). Springer.

Bayoudh, M., Travé-Massuyès, L., & Olive, X. (2008). Coupling continuous and discrete event system techniques for hybrid system diagnosability analysis. In *18th european conf. on artificial intel.* (pp. 219–223).

Bayoudh, M., Travé-Massuyès, L., & Olive, X. (2009). Diagnosis of a Class of Non Linear Hybrid Systems by On-line Instantiation of Parameterized Analytical Redundancy Relations. In *20th international workshop on principles of diagnosis* (p. 283-289).

Benazera, E., & Travé-Massuyès, L. (2009, October). Set-theoretic estimation of hybrid system configurations. *Trans. Sys. Man Cyber. Part B*, *39*, 1277–1291. doi: 10.1109/TSMCB.2009.2015280

Bregon, A., Daigle, M., Roychoudhury, I., Biswas, G., Koutsoukos, X., & Pulido, B. (2014, May). An event-based distributed diagnosis framework using structural model decomposition. *Artificial Intelligence*, *210*, 1-35.

Bregon, A., Narasimhan, S., Roychoudhury, I., Daigle, M., & Pulido, B. (2013, October). An efficient model-based diagnosis engine for hybrid systems using structural model decomposition. In *Proceedings of the annual conference of the prognostics and health management society, 2013.*

Cocquempot, V., El Mezyani, T., & Staroswiecki, M. (2004, July). Fault detection and isolation for hybrid systems using structured parity residuals. In *5th asian control conference* (Vol. 2, p. 1204-1212). doi: 10.1109/ASCC.2004.185027

Daigle, M. (2008). *A qualitative event-based approach to fault diagnosis of hybrid systems* (Unpublished doc-

toral dissertation). Vanderbilt University.

Daigle, M., Bregon, A., & Roychoudhury, I. (2015, September). A Structural Model Decomposition Framework for Hybrid Systems Diagnosis. In *Proceedings of the 26nd international workshop on principles of diagnosis.* Paris, France.

Daigle, M., Koutsoukos, X., & Biswas, G. (2009, July). A qualitative event-based approach to continuous systems diagnosis. *IEEE Transactions on Control Systems Technology*, *17*(4), 780–793.

Daigle, M., Roychoudhury, I., Biswas, G., Koutsoukos, X., Patterson-Hine, A., & Poll, S. (2010, September). A comprehensive diagnosis methodology for complex hybrid systems: A case study on spacecraft power distribution systems. *IEEE Transactions of Systems, Man, and Cybernetics, Part A*, *4*(5), 917–931.

Daigle, M., Roychoudhury, I., & Bregon, A. (2014, September). Qualitative event-based fault isolation under uncertain observations. In *Annual conference of the prognostics and health management society 2014* (p. 347-355).

Daigle, M., Roychoudhury, I., & Bregon, A. (2015). Qualitative event-based diagnosis applied to a spacecraft electrical power distribution system. *Control Engineering Practice*, *38*, 75 - 91. doi: http://dx.doi.org/10.1016/j.conengprac.2015.01.007

Gaudel, Q., Chanthery, E., & Ribot, P. (2015). Hybrid particle petri nets for systems health monitoring under uncertainty. *International Journal of Prognostics and Health Management.*, *6*.

Henzinger, T. A. (2000). *The theory of hybrid automata*. Springer.

Hofbaur, M., & Williams, B. (2004). Hybrid estimation of complex systems. *IEEE Trans. on Sys., Man, and Cyber, Part B: Cyber.*, *34*(5), 2178-2191. doi: 10.1109/TSMCB.2004.835009

Koutsoukos, X., Kurien, J., & Zhao, F. (2003). Estimation of distributed hybrid systems using particle filtering methods. In *In hybrid systems: Computation and control (hscc 2003). springer verlag lecture notes on computer science* (pp. 298–313). Springer.

Mosterman, P., & Biswas, G. (2000). A comprehensive methodology for building hybrid models of physical systems. *Artificial Intel.*, *121*(1-2), 171 - 209. doi: DOI: 10.1016/S0004-3702(00)00032-1

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

Narasimhan, S., & Biswas, G. (2007, May). Model-Based Diagnosis of Hybrid Systems. *IEEE Trans. Syst. Man. Cy. Part A*, *37*(3), 348-361.

Narasimhan, S., & Brownston, L. (2007, May). HyDE: A General Framework for Stochastic and Hybrid Model-based Diagnosis. In *Proc. of the 18th int. ws. on principles of diagnosis* (p. 186-193).

Reiter, R. (1987). A Theory of Diagnosis from First Principles. *Artificial Intelligence*, *32*, 57-95.

Rienmüller, T., Bayoudh, M., Hofbaur, M., & Travé-Massuyès, L. (2009). Hybrid Estimation through Synergic Mode-Set Focusing. In *7th ifac symposium on fault detection, supervision and safety of technical processes* (p. 1480-1485). Barcelona, Spain.

Roychoudhury, I., Daigle, M., Bregon, A., & Pulido, B. (2013, March). A structural model decomposition framework for systems health management. In *Proceedings of the 2013 IEEE aerospace conference.*

Trave-Massuyes, L., & Pons, R. (1997). Causal ordering for multiple mode systems. In *Proceedings of the eleventh international workshop on qualitative reasoning* (pp. 203–214).