Online Model-based Diagnosis to support autonomous operation of an Advanced Life Support System

Gautam Biswas, Eric-J. Manders, John Ramirez Nagabhusan Mahadevan, and Sherif Abdelwahed

The authors are are with the Department of Electrical Engineering and Computer Science, and the Institute for Software Integrated Systems, Vanderbilt University, P.O. Box 1824 Station B, Nashville, TN 37235-1592 USA. Phone: +1 615 322-0732, fax: +1 615 343-6702. (e-mail: {gautam.biswas}@vanderbilt.edu)

Online Model-based Diagnosis to support autonomous operation of an Advanced Life Support System

Gautam Biswas, Eric-J. Manders, John Ramirez Nagabhusan Mahadevan, and Sherif Abdelwahed

Abstract

This paper describes methods for online model-based diagnosis of subsystems of the Advanced Life Support System (ALS). The diagnosis methodology is tailored to detect, isolate, and identify faults in components of the system quickly so that fault-adaptive control techniques may be applied to maintain system operation without interruption. We describe the components of our hybrid modeling scheme and the diagnosis methodology, and then demonstrate the effectiveness of this methodology by building a detailed model of the Reverse Osmosis (RO) system of the Water Recovery System (WRS) of the ALS. This model is validated with real data collected from an experimental testbed at NASA JSC. A number of diagnosis experiments run on simulated faulty data are presented and the results are discussed.

Index Terms

Fault diagnosis, detection, estimation, Advanced Life Support systems, Water Recovery System, Reverse Osmosis system.

The authors are are with the Department of Electrical Engineering and Computer Science, and the Institute for Software Integrated Systems, Vanderbilt University, P.O. Box 1824 Station B, Nashville, TN 37235-1592 USA. Phone: +1 615 322-0732, fax: +1 615 343-6702. (e-mail: {gautam.biswas}@vanderbilt.edu)

I. INTRODUCTION

The support of human life for extended durations in the hostile environment of space critically depends on a set of complex technical systems that contain or interact with biological and chemical processes. The NASA Advanced Life Support Systems (ALS) program, a component of the larger Advanced Human Support Technology (AHST) Program, was created to explore new technologies to support future manned missions in space [1]. The exploration program for the agency has also recently gained a new focus toward human space exploration outside of low earth orbit [2]. Extended duration missions include a Lunar base, a human mission to Mars, and the permanent occupation of the International Space Station (ISS). The ALS is a mission critical subsystem of such missions that must operate at a high level of autonomy, so as not to detract from the mission specific tasks of the crew. While this places high reliability of the individual components of the ALS the system as a whole needs to be designed for robust operation under a range of possible scenarios. Ideally, the integrated system will have the ability to adapt to changing mission objectives, crew configurations, and respond to unplanned events.

Achieving good operational performance is critically dependent on the ability to monitor and analyze the operation of the physical system, and, when necessary, respond by retuning or reconfiguring the controllers or the system so that important functionalities are not lost or degraded. The overall objective is to develop techniques so that the system can *adapt* to faults by reconfiguring itself and/or its controllers. Such adaptivity can be achieved either in closed loop or operator assisted (human in the loop) schemes. Adaptivity is obtained through process diagnosis capabilities, i.e., systematic ways to explain deviations from nominal system behavior in terms of failures and degradations in system components. Diagnosis combines fault detection, fault isolation, and fault identification tasks [3].

An ALS comprises multiple coupled sub-systems, such as (*i*) a Water Recovery System (WRS), (*ii*) an Air Revitalization System (ARS), (*iii*) a Power generation system, (*iv*) a Thermal control system, (*v*) a Biomass production system, (*vi*) a Food production sub-system, and (*vii*) a Solid waste collection and conditioning system. Analysis and control of the system is challenging due to the multiple interacting control loops [4]. These loops, for fluid flow, energy management, thermal control, bio-regeneration, gas transfer, and chemical production cover multiple physical (energetic) domains and operating regimes, and operate at multiple time scales. An effective way to describe the behavior of the controlled physical sub-systems is to model them as hybrid dynamical systems, which capture both the both continuous and discrete dynamics [5].

A wide array of techniques have been developed for monitoring and diagnosis of physical systems. The complexity of these systems and the increasing need for robustness and reliability over a wide range of operating modes has made *model-based* approaches to fault detection and isolation (FDI) schemes increasingly attractive. Model-based schemes use a small number of system measurements and exploit analytical redundancy relations in the model to analyze discrepancies in the observed system behavior. This provides for cheaper and more reliable diagnosis solutions than what may be achieved by classic hardware redundancy techniques, or methods that are designed for individual fault by fault analysis. However, the increasing complexity of the physical systems also implies that the FDI system designer must contend with uncertainties in the model structure and its parameter values, as well as noise in the measurement data. For hybrid dynamical systems we require an on-line scheme where the objective is to monitor system behavior through a succession of operating modes, detect discrepancies in behavior when they occur, and isolate them in a timely manner.

We have developed a comprehensive approach, *Fault Adaptive Control Technology* (FACT), that encapsulates modeling the physical systems as hybrid dynamical system models, tracking dynamic behavior on-line, performing model based FDI when discrepancies are detected between expected and observed behavior, and applying control reconfiguration techniques to accommodate for the faulty situation [6]. In

this paper, we focus on the problem of model-based FDI, and its application to the ALS system.

Faults can typically occur in process components, sensors, actuators, and the controller. This work focuses on the detection and isolation of faults in process components. These faults directly affect the process dynamics and, therefore, are considered to be *multiplicative faults*. In analytic model-based approaches [7], [3], these faults are linked to coefficients in the state space matrices or the transfer functions of the system. Given that our goal for fault isolation is to implicate system components, we have adopted a modeling paradigm based on bond graphs [8] that provides a direct correspondence between inferred faults expressed as parameter value changes in the model, and the actual physical system components. This is important if the diagnosis results are required for maintenance and fault-adaptive control tasks.

Faults cause deviations in the dynamic behavior of the system. In our work we define a fault mathematically, as a discrete and persistent change in a model parameter. This is the notion of *abrupt faults*, which corresponds to a very quick change in component's performance. The change in performance can be attributed to a breakage or a degradation in the component. A contrasting notion is an *incipient fault*, which correspond to a gradual drift in a component parameter value [9]. An abrupt fault in a component parameter causes a transient response in the system variables, and this can be attributed to the system attempting to correct the energy imbalance introduced by the fault, i.e., the change in the component parameter value. The transient behavior vanishes after an interval, and for faults in certain components evidence of the fault may not be observable in the measured variables after the transient disappears. For these faults, the transient response presents a finite window of opportunity to perform FDI.

Our FDI scheme, named TRANSCEND, is based on the analysis of the fault transient after the occurrence of an abrupt fault. A unique aspect of TRANSCEND is that it combines tools and techniques from statistical signal processing for fault detection and symbol generation from the residual signals [10] with computational intelligence techniques for qualitative fault isolation [11]. Analytic redundancy relations (ARR) in the model are converted to a symbolic form that captures the transient dynamics, and constraint analysis techniques [12] form the basis of efficient fault isolation algorithms. Further, the use of a qualitative framework implicitly accommodates model uncertainty. On the other hand, data uncertainty is handled by statistical techniques for fault detection and parameter estimation in this framework.

Section 2 introduces the basic building blocks of our model-based scheme for FDI of process components, emphasizing the importance of component-based modeling and the analysis of transient behavior in continuous systems. Section 3 describes the Reverse Osmosis (RO) sub-system of the Water Recovery System (WRS) in a physical ALS test-bed that is the application focus of this paper. We describe in detail how the system is modeled for diagnosis purposes. Section 4 describes the results of our experiments using the FACT tools for the RO system, and section 5 presents the conclusions of this work.

II. MODEL-BASED FDI OF ABRUPT FAULTS

In this section, we develop our approach for FDI of abrupt faults from transient behavior. We describe the modeling paradigm, the models used for diagnosis, and the basic components of our model-based FDI scheme. Details of each of these topics have been discussed in our earlier papers [6], [11].

A. Principles of model-based FDI

Building models at the right level of detail is a critical first step in the development of a model-based FDI scheme. The choice of the model representation and the level of detail included in the model determines the set of faults that are linked to model parameters, and, therefore, can be distinguished during fault isolation.

A.1 Modeling with bond graphs

Bond graphs (BG) define a domain-independent topological modeling language that captures energybased interactions among the different physical processes that constitute a dynamic system [8]. A bond graph model can be derived in a systematic manner from a given system description [13]. Physical system components and sub-systems are modeled by a combination of one or more generic physical processes, and represented using capacitors, inertias, dissipators, transformers, gyrators and sources. Each of the processes can capture linear and nonlinear behaviors. Capacitor and inertia elements capture energy storage processes in the system, dissipator elements model mechanisms by which the system dissipates energy to the environment, sources represent the interaction of the system with it's environment (energy flow in and out of the system), and transformers and gyrators are mechanisms that transform energy from one form to another. These elements form the vertices of the bond graph model. Additional vertices impose conservation of energy laws at idealized connecting points between the elements. These vertices, referred to as 0and 1-junctions, are domain independent generalizations of Kirchoff's laws. The energy pathways through which subsystems/processes exchange energy are represented as the graph edges, called bonds. Each bond is associated with two energy-related variables: effort (e) and flow (f), where $e \cdot f$ equals power, the rate of energy exchange among the two components connected by the bond. Other edges, called signals, model information flow between systems or components.

A component fault manifests as a change in the value of a bond-graph element parameter. These faults affect the coefficients of the system matrix in the state-space or transfer function representation of the system, that is, they have *multiplicative* effects on the system dynamics. The bond graph representation preserves a one-to-one correspondence between the component parameters and the physical components of the system. As a contrast, in state-space or transfer function representations the model coefficients are typically functions of more than one physical component parameter. For fault isolation this implies that the bond graph representation creates a direct link between the change in a parameter value with a fault in a specific system component.

A.2 Abrupt Faults

The abrupt fault profile is widely used in the FDI literature [14] as a mathematical representation of temporal change in a faulty parameter value. It is often applied as a modeling abstraction when the real fault profile is unknown[15]. This allows for the straightforward definition of decision functions for fault detection that can be evaluated in a statistical hypothesis testing framework [14].

In our work, we adopt a strict mathematical interpretation of an abrupt fault. An abrupt fault is a component parameter change that occurs at a much faster rate than the nominal dynamics of the system. In a sampled data system, the sampling rate is tuned to the nominal dynamics of the system. Therefore, the fault manifests as an instantaneous jump in a parameter value, and system response to the this change is a transient that is superimposed on the continuous dynamic behavior. We define a *fault transient* as the system behavior in response to an abrupt component fault.

B. The TRANSCEND approach

Fig. 1 shows a block-diagram of the FACT architecture that includes the TRANSCEND FDI scheme. TRANSCEND follows the residual generation and evaluation scheme for model-based FDI [9]. Residual generation $(r = y - \hat{y})$ is realized through robust tracking of nominal system behavior using extended Kalman filter (EKF) techniques [16]. Residual evaluation includes statistical fault detection and symbol generation techniques, and a novel fault isolation method that is based on the qualitative analysis of the system dynamics immediately after the time point of fault occurrence. This results in a reduced fault hypothesis set, and quantitative parameter estimation techniques are applied to the reduced set to uniquely isolate and identify the fault [17]. The extension of these methods to hybrid systems complicates the analysis in that discrete mode changes, and, therefore, model switches occur while tracking and analyzing system behavior [5]. An automaton model is employed to switch system models when mode changes occur [6]. The rest of this section outlines the main components of the TRANSCEND system.



Fig. 1. FACT architecture with TRANSCEND FDI components.

B.1 Qualitative fault isolation from fault transients

The core component of TRANSCEND is the qualitative fault isolation engine that performs symbolic analysis of the fault transient to generate and refine fault hypotheses. The model representation used for this analysis is a *Temporal Causal Graph* (TCG) that captures the relevant causal and dynamic relations between the system variables [11]. The TCG is in many ways similar to the signal flow graph (SFG) representation of a system [10], and is derived systematically from a bond graph model. Edges define relations between variables (nodes) in the TCG. These relations can be algebraic (direct (+) or inverse (-) proportionality, and equivalence ('=')), or they can define a temporal relation (using a specifier 'dt' implying that the successor node has an integral relation with the predecessor vertex value). The TRANSCEND algorithms exploit the properties of the TCG that lend themselves to symbolic processing of system dynamics using graph propagation algorithms. The TCG captures the analytical redundancy relations between variables while retaining system topology from the bond graph model.

The fault isolation engine follows the generate-and-test approach to residual evaluation using the TCG structure. Qualitative transient behavior is expressed as a *fault signature* that describes the expected fault transient *immediately after fault occurrence*. The signature corresponds to a qualitative interpretation of the Taylor series expansion of the residual around the time point of fault occurrence [18]. The order of the signature is defined by the highest derivative computed (a design parameter). Symbolic values for the elements of a signature are: '+' for a positive or increasing value, '0' for a zero or unchanged value, and '-' for a negative or decreasing value. An unknown value is represented by '.'. The description of a

fault signature in terms of the behavior around the point of fault occurrence is unique to the TRANSCEND approach.

Fault isolation is triggered by the first non-zero magnitude symbol that is output by the signal-to-symbol generation module. This initial symbol reflects the magnitude deviation in the residual at the onset of the fault transient. The *hypothesis generation* step produces a set of fault hypotheses that can explain the observed deviation. A fault hypothesis consists of a candidate parameter with a direction of change for the parameter value and a fault signature for each of the measured variables. During *hypothesis refinement*, the signatures are compared with symbolic values computed from the measurements using a scheme called *progressive monitoring*. When a match fails, the candidate is dropped. Further details of the qualitative fault isolation scheme are presented in [11]. Extensions of this scheme to diagnosis of hybrid systems appears in [17].

B.2 Signal-to symbol transformation

A critical aspect of any qualitative analysis scheme for continuous dynamic systems involves the computation of symbolic information that change over time. The mapping is essentially an estimation problem, where the local dynamics are mapped onto the qualitative space, +, 0, and -. The signal generation process is closely linked to the detection process, especially if we consider that the first non-zero symbol corresponds to the detection of the fault. The details of the scheme have been reported on in [17], [11].

An alternative approach is described in [10], where the detection of the transient is decoupled from the symbol generation task. By developing a dedicated transient detection scheme, a much higher sensitivity can be obtained, at a cost of a more complex solution. We can show that with increasing fault size, that the precision of the fault isolation scheme improves, as more symbols can be robustly determined from the residual data.

B.3 Parameter estimation

The hybrid and possibly non-linear system behavior makes it hard to apply traditional parameter estimation techniques. TRANSCEND utilizes a novel mixed simulation-and-search algorithm [17] to estimate the value of the faulty parameter. Under the single fault assumption the identification problem is simplified to a one dimensional parameter optimization problem. Starting with the results of the qualitative fault isolation scheme that returns a set of potential fault hypotheses, we run multiple optimizations, and each one estimates one fault parameter value. The parameter estimation scheme uses data collected from the time point of fault detection, the current state variable values, and a set of N measurement samples that includes the system input and output signals. The estimation scheme is based on an optimization algorithm (technically any non linear optimization algorithm may be employed), and it finds the fault parameter value that minimizes the least square error between the expected system output generated by the simulator and the available measurement values over the N samples using a greedy search algorithm. The simulator uses the hybrid automaton model of the system, derived from our hybrid bond graph models.

III. DIAGNOSIS OF ALS SUBSYSTEMS

We demonstrate TRANSCEND's diagnosis scheme by applying it to the Reverse Osmosis (RO) subsystem of the Advanced Water Recovery System (AWRS). A real AWRS testbed was designed and built at the NASA Johnson Space Center (JSC) [19], [20]. Fig. 2 shows a block diagram of the system with its four main components. Waste water first enters the Biological Water Processor (BWP) subsystem, which removes organic matter and ammonia from the water. The effluent from the BWP is fed into the Reverse Osmosis (RO) subsystem, which uses a membrane system to remove inorganic matter and particles from the water. The combination of the BWP and RO can typically clean about 85% of the input waste water. The remaining 15%, which is concentrated brine formed in the RO is passed into the Air Evaporation System (AES), which recovers additional water through a cyclic evaporation and condensation processes. The fourth component implements a post processing step, where ultraviolet treatment is applied to the output of the RO and the AES to remove trace contaminants and generate potable water.



Fig. 2. Block diagram for the Advanced Water Recovery System (AWRS).

We generated a model of the RO system of the above AWRS for our diagnosis studies. The first step was to construct a HBG model of the subsystem, and then real data collected from the system was used to estimate the HBG model parameters. Details of the model are presented in this section.

A. The Reverse Osmosis (RO) sub-system

We start with a brief description of the RO subsystem components [19], [20], and then describe the model that we have developed.

A.1 Principle of operation

Fig. 3 shows the RO subsystem process diagram. The inputs to and outputs from the subsystem and the measured variables are clearly marked. The key component of the RO is a membrane [20]. Input water from the BWP is pushed at high pressure through the membrane. Clean water (permeate) leaves the system, and the remaining water (with a larger concentration of brine) is recirculated in a feedback loop. As a result, the concentration of impurities in the recirculating water increases with time.

The system cycles through three operating modes, which are set by the 4-way multi-position valve. The feed pump, which is on in all modes, pulls effluent from the BWP and creates a flow into the system through a coiled pipe, which acts as a tubular reservoir. In the *primary* mode (valve setting 1), the input flow is mixed with the water in the primary recirculation loop. The recirculation pump boosts the liquid pressure as it flows into the membrane. The flow through causes dirt to accumulate in the membrane, which increases the resistance to the flow through it, thus causing the outflow from the system to decrease with time. The testbed at JSC was setup to transition to the *secondary* mode of operation after a predetermined time interval. In secondary mode (valve setting 2), the recirculating fluid is routed back to the membrane in a smaller secondary loop. This causes the liquid velocity (and, therefore, flowrate) to increase, and as a result the outflow from the system does not keep decreasing as sharply as it does in the primary loop. As clean water leaves the system, the concentration plus the collection of impurities in the membrane decreases its performance sufficiently, and again after a predetermined time interval, a transition is made to the *purge* mode (valve setting 3), where the recirculation pump is turned off, and concentrated brine is pushed out to the AES subsystem. Following the purge operation, the system goes back to primary mode.



Fig. 3. High level process engineering diagram for the RO system.

In real operations, the system goes through a mechanical cleaning (the sloughing mode) once every 5-6 cycles to prevent complete clogging of the membrane. The sloughing mode is not explicitly modeled in this paper. Overall, systems engineers determined that the primary plus secondary loop operations cleaned about 85% of the water, and the remaining 15% was sent to the AES for processing.

The JSC test bed was extensively instrumented. For our work, we have used five of the measurements (see Fig. 3): (i) the pressure immediately after the recirculation pump, P_{pump} , (ii) the pressure of the permeate at the membrane, P_{memb} , (iii) the pressure of the liquid in the return path of the recirculation loop, P_{back} , (iv) the flow of the effluent, F_{perm} , and the conductivity of liquid in the return path of the recirculation system. Shown are from top to bottom the membrane pressure, backflow pressure, and conductivity. The pump pressure in the actual system was not available (a differential pressure over the pump was measured), and the actual permeate flow measurement is sparse and unreliable. The mode transition sequence was fixed at design time, and the changes from primary (M1) to secondary (M2) to purge are shown as control signals.

A.2 Modeling for diagnosis

The HBG model for the RO subsystem was developed using the system schematics and by developing a good understanding of the behavior of each of the components. The model captures three principal physical domains in which the system operates. The *mechanical* and *hydraulic* domains, which cover the pump and fluid flow, are the primary energy domains. In addition, to take into account the effects of impurities in the water on the flow process, and the fact that these impurities are time-varying, we explicitly model the fluid *conductivity* domain and its interactions with the flow process using bond graph elements. For the diagnosis model, we have abstracted the energy interaction between the mechanical and the hydraulic domains by a single pump efficiency parameter. As a simplification, other interactions between the mechanical parts of the pump and the fluid conductivity are not modeled. A primary innovation in this model is the ability to capture the interaction between the hydraulic and conductivity domains in the bond graph using modulating signals.

Fig. 5 shows the flattened HBG model, without component and subsystem hierarchies. Switching elements in the model capture the three modes of operation, and the switching function into each mode is



Fig. 4. Nominal operation for the RO system.

defined by a logical expression. The three model domains are described in greater detail below.

The Mechanical Domain

The mechanical domain model includes the feed pump and the recirculation pump. The feed pump motor is modeled as an flow source, Se_{fp} . The motor creates a momentum in the pump's rotor system, whose mass is modeled by the inertia I_{fp} . The dissipation of energy due to friction in the rotation mechanism is modeled by the resistance R_{fp} . The conversion of rotational speed to fluid flow rate (i.e., the conversion of mechanical to fluid energy) is captured by a transformer TF. The recirculation pump motor is modeled as an effort source Se_{ep} , with a pump rotor with inertia, I_{ep} . The mechanical frictional losses are modeled by a resistance, R_{ep} . The conversion of mechanical energy to fluid flow is captured in the gyrator, GY. In the purge mode, the 1-junction connecting the source effort to the rest of the system is switched off, and this effectively disconnects the pump from the system. The fluid pressure in the loop at the output of the pump, P_{pump} is a measured variable (effort e37 on the energy bond that connects the pump to the hydraulic domain).

The Hydraulic Domain

The hydraulic domain contains four main components: (*i*) the tubular reservoir,(*ii*) the back-flow loop, (*iii*) the membrane module, and (*iv*) the pipe connecting the reservoir to the membrane. A drain pipe component is active only in the purge mode. The fluid is assumed to be incompressible but the pressure flow-rate relationship in the pipes is nonlinear [21]. The reservoir is modeled as a capacitor, C_c connected to a 0-junction. In reality, C_c represents the combined capacitance of the coil and the adjacent pipes that are active in the the primary and secondary loop configurations. The pipe connecting the coil to



Fig. 5. Hybrid Bond Graph model of the RO system. Components that are not active in primary mode are shaded gray.

the membrane is modeled as a resistance, R_{pipe} . The membrane assembly is modeled using a variable resistance parameter, R_{memb} , connected in parallel to a capacitor, C_{memb} . The capacitor represents the fluid that collects in the membrane component, and the state variable associated with C_{memb} defines the fluid pressure at the membrane, i.e., P_{memb} , a measured variable (effort e16). R_{memb} models resistance to flow through the membrane. This parameter is modulated by the conductivity of the water flowing in the system, K, to capture the increase in its value as time progresses¹. The flow through the membrane, F_{perm} , is a measured flow variable, f 25. The water that does not pass through the membrane has a greater concentration of impurities and is returned to the reservoir. This is modeled as the back flow loop, which is active in both the primary and secondary modes of operation. The pressure in the return path of the loop, P_{back} is measured, and corresponds to effort variable e1. The pipe in this loop is modeled as a resistance R_{brine} . The change in conductivity, modeled in the conductivity domain, is a function of the flow-rate variable, f 15 in the back flow loop. In the purge mode, the resistor R_{drain} represents the the pipe to the AES. The switch to this mode also results in the back flow pipe resistance being increased to a very high value, R_{valve2} , and the higher resistance prevents the flow back into the loop, therefore, most of the brine is purged into the AES.

The Conductivity Domain

The conductivity of the water, K, is a measure of the concentration of impurities in the water, and corresponds to the effort variable e^{35} . The value of K is time varying and in the bond graph it is captured as the state variable for element C_K . The state of C_K depends on the value of its inputs ($S_{f:1}$ or $S_{f:2}$),

¹ The resistance of the membrane is reset after the purge cycle because the slough mode is not taken into account in the model. The initial value of this resistance is increased by a small amount at the beginning of each primary loop to improve the correspondence to the real data

the value of R_K and the state of C_{K2} . C_{K2} is used to capture the effect fluid levels in the system have on K, and it is essential for properly modeling the behavior of K in the purge mode. The inputs affecting K in the primary and secondary mode of operation are $S_{f:1}$ and $S_{f:2}$, respectively. Both of these flow sources are modulated by two system dependent variables, the integral of the flow rate into the backflow loop (C_{brine}) and the flow rate coming into the membrane f_3 . $S_{f:1}$ and $S_{f:2}$ differ in that their modulation functions, f_3 and f_2 , respectively, assign different weights to the modulating variables. This allows us to model the effect of switching from primary to the secondary mode as an increase in the source flow going into C_K . In the purge mode, both $S_{f:1}$ and $S_{f:2}$ become zero, and the value of K, the state variable of C_K is maintained as the state of C_{K2} decreases substantially as the concentrated brine is purged into the AES. This is implemented using a resistance R_{purge} connected in parallel to C_{K2} , causing the capacitor to drain through the resistance.

B. Estimating Model Parameters

Parameter values for the model are estimated using data from the JSC testbed (Fig. 4). To achieve this, parts of the RO system were isolated, the appropriate equation forms describing the parameter and measured value relations were derived from the bond graph, and numerical value of the parameters obtained through least squares estimation techniques. Table I lists the estimated nominal parameter values. The model was subsequently validated through simulation experiments.

| Param. | unit (SI) | value |
|-------------|-----------------------|-------|
| C_{k} | $ml/min \cdot mS$ | 565 |
| C_c | m^5/N | 1.5 |
| C_{brine} | m^5/N | 8 |
| R_{brine} | $N \cdot / m^5$ | 220 |
| C_{memb} | m^5/N | 0.6 |
| R_{memb} | $N\cdot/m^5$ | 26.0 |
| I_{ep} | $N \cdot m \cdot s^2$ | 2 |
| R_{ep} | $N \cdot s \cdot m$ | 0.1 |
| I_{fp} | $N \cdot m \cdot s^2$ | 0.1 |
| R_{fp} | $N\cdot/m^5$ | 0.1 |
| R_{pipe} | $N\cdot/m^5$ | 69.0 |

 TABLE I

 Nominal values for the RO system bond graph parameters

C. Implementation of the HBG Model using the FACT Tool Suite

The model for this system was created using the *Generic Modeling Environment* (GME) [22], developed at the Institute for Software Integrated Systems (ISIS) at Vanderbilt University. The meta-programmable modeling environment is configured using the FACT paradigm [6]. This includes a component based, hierarchical HBG modeling environment, model interpreters, and an execution environment that implements the FACT architecture described in section 2. Fig. 6 shows the RO system model as realized using the HBG modeling tool. The individual RO system components described in the last section appears as blocks in the top-level model. Each component is in turn modeled as a HBG model fragment, and the model fragment for the recirculating pump is shown as an example.

The FACT paradigm has two model interpreters. One interpreter generates a MATLAB/SimulinkTM simulation model, that also supports the simulation of fault scenarios for experiments. A second interpreter creates a flattened HBG model representation that is exported to a model database. The run-time FACT



Fig. 6. Top level of the hierarchical model for the RO system created using the hybrid bond graph modeling tool in GME. The effort pump is shown at the next level down the hierarchy (flat bond graph model).

system uses this database to generate (*i*) the state space models used by the EKF in the hybrid observer and the optimization algorithm during parameter estimation, and (*ii*) the TCG models (for each mode) used by the qualitative fault diagnosis algorithms. Fig. 7 shows the TCG for the model in the primary operating mode.

IV. EXPERIMENTS AND RESULTS

We demonstrate the effectiveness of the TRANSCEND scheme by running simulation experiments on a number of fault scenarios. The system was simulated for three cycles of operation. The sample time in the simulation is an order of magnitude faster than the sample time in the real data, so we can consider the simulation to be 'oversampled'. Although empirical information on sensor noise is not available, measurement noise was simulated as Gaussian white noise with a noise power level set at 2% of the average signal power for each measurement. Fault scenarios were created that correspond to abrupt faults in the pump (loss of efficiency and increased friction in the bearings), membrane (clogging), and the connecting pipes (blocks).

We illustrate the operation of TRANSCEND with the detailed results for one scenario, a decrease in



Fig. 7. Temporal Causal Graph corresponding to the hybrid bond-graph in Fig. 5 for the primary operating mode.

recirculation pump efficiency. Subsequently, we demonstrate the generality of the approach by evaluating the results on four other scenarios.

A. Analysis of fault scenario: Decrease in pump efficiency

A decrease in the recirculating pump efficiency is modeled by a decrease in the value of a pump component parameter, the gyrator, GY, indicated as GY^- . Fig. 8 shows the plant data including the controller signals and the output of the observer, and the computed residual signals. The fault is introduced in the second operating cycle, while the system is in the primary operating mode.

Fig. 9 illustrates the steps in the qualitative fault isolation scheme for the GY^- scenario. Each of the four steps that happen in sequence over a period of time, corresponds to an event when a new non-zero symbol is generated, reflecting a qualitative change in the transient dynamics of the fault response. Step 0 is defined as the time when the fault is first detected, i.e., the first statistically significant non-zero residual is observed. The time at which this event occurs is marked as the difference between the current time and the time of fault occurrence, $t - t_f$. $t - t_f$ measures the delay in fault detection after the actual fault occurrence.

At each step, the current set of derived hypotheses is displayed in tabular form. The symbolic residual value, indicating qualitative magnitude and slope for each measurement, is shown in the block marked 'actual' at the top. The '+' and '-' symbols correspond to an observed positive (negative) value in the magnitude field of the residual signal, and an positive or negative slope value in the slope field of the signal. A '0 'implies no change, and a '' value implies an unknown value. The predicted signatures are



Fig. 8. Plant data (*l*) and residual (*r*) for fault GY^- with fault size 5%, occurring at $t_f = 17500$ (*s*) = 5(*hours*) (primary mode in the second cycle).

grouped by fault candidate and shown for all measurements. For this system, we found it sufficient to work with fourth order signatures. Other work discusses the tradeoff between using higher order signatures and their discriminating power [18].

At step 0 we observe that the residual for the pump pressure deviates in the negative direction $(e37 = (-, \cdot))$. The hypothesis generation step that is triggered by this event results in twelve hypothesized fault candidates. The next observed change, the symbolic event at step 1, occurs when the flow-rate through the membrane also shows a negative deviation $(f25 = (-, \cdot))$. Fault hypotheses whose signatures are not consistent with this (symbolic) residual value are dropped, and the fault set is reduced to seven hypotheses. At step 2, the deviation in the back flow pressure, e1, becomes significant, and at step 3 the deviation in the conductivity, e35, becomes significant also. When symbolic information can not discriminate between the remaining fault hypotheses, qualitative fault isolation terminates. Three candidates, GY^- , I_{ep}^+ , and R_{ep}^+ remain after the qualitative fault isolation completes. In addition to the actual (true) fault candidate, these correspond to an increase in pump inertia (unlikely fault), and an increase in the pump resistance (e.g., increase in friction in the pump bearings), respectively.

Further refinement in the fault set can only be done by quantitative analysis. After sufficient number of samples are collected (this is currently fixed to 200), the parameter estimation scheme is initiated. Results

| Step 0 ($t - t$ | $t_f = 200(s))$ | Step 1 (t - | $t_f = 880(s))$ |
|-----------------------|---|-------------------|-------------------|
| Actual | e1: 0 0 | Actual | e1: 0 0 |
| | e16:0 0 | | e16:0 0 |
| | e35:0 0 | | e35:0 0 |
| | e37: | | e37:- 0 |
| | 125:0 0 | | f25: |
| GY^- | e1: $0 - \cdot \cdot \cdot$ | GY^{-} | e1: 0 – · |
| | e16:0 - · · · | | e16:0 - · |
| | $e_{35:0} - + \cdots$ | | e35:0 - + |
| | $e_{3/2} + - \cdot \cdot$ | | $e_{3/2} + -$ |
| <i>T</i> + | 125:0 | 7+ | 21: 0 |
| 1 ep | e1. 0 = · · · | 1 ep | e1. 0 = · |
| | $e_{35:0} = +$ | | $e_{35:0} = +$ |
| | e37:-+-·· | | e37:-+- |
| | f25: 0 | | f25: 0 |
| R^+_{am} | e1: 0 0 - · · | R^+_{am} | e1: 0 0 - |
| ep | e16:0 0 - · · | ep | e16:0 0 - |
| | $e35:0 \ 0 - + \cdot$ | | e35:0 0 - |
| | $e37:0 - + - \cdot$ | | e37:0 – + |
| | f25:00 ···· | | f25:00 · |
| R_nipe | e1: $0 + \cdot \cdot \cdot$ | R _{bmin} | e1: 0 - + |
| P.P.C | $e16:0 + \cdot \cdot \cdot$ | orine | e16:0 - + |
| | e35:0 + - · · | | e35:0 0 + |
| | e37:0 - + | | e37:0 0 – |
| | f25: 0 · · · · | | f25: 0 – · |
| C^+_{memb} | e1: – + \cdot \cdot | I_{fp}^{-} | e1: 0 - + |
| | e16:-+ | | e16:0 0 + |
| | $e_{35:0} + - \cdot \cdot$ | | $e35:0 0 + \cdot$ |
| | $e_{37:0} - + \cdot \cdot$ | | e37:0 0 - · |
| <u>a</u> = | 125: | | f25: 0 0 · |
| C_c | $e_1: - + \cdots$ | R_{fp} | el: 0 0 - |
| | $e_{10.0} + \cdots$ | | el6:0 0 0 - |
| | $e_{37:0} = \pm$ | | e35:0 0 0 - |
| | f25: 0 | | £25:000 |
| R_1^- . | e1: $0 - + \cdot \cdot$ | TF^+ | e1: 0 - + |
| orine | $e16:0 - + \cdot \cdot$ | | e16:0 0 + |
| | e35:0 0 + - \cdot | | e35:0 0 + |
| | $e_{37:0} 0 - + \cdot$ | | e37:0 0 – |
| | f25: 0 – · · · | | f25:00 · |
| R_{memb}^{-} | e1: $0 - + \cdot \cdot$ | | |
| | $e16:0 - + \cdot \cdot$ | | |
| | $e_{35:0} 0 + - \cdot$ | | |
| | $e_{3/10} = 0 = + +$ | | |
| <i>c</i> ⁺ | 123:+ - | | |
| C_k | $e_1: 0 - + \cdot \cdot \\ e_1 = e_1 + \cdot \\$ | | |
| | $e_{35} = 0 \pm -$ | | |
| | $e_{37:0} 0 - +$ | | |
| | f25:+ - | | |
| 1. | e1: $0 - + \cdot \cdot$ | | |
| J P | $e16:0 0 + \cdots$ | | |
| | $e_{35:0} 0 + - \cdot$ | | |
| | e37:0 0 - + · | | |
| | f25: 0 0 · · · | | |
| R_{fn}^{-} | e1: 0 0 - + · | | |
| JP | e16:0 0 0 + | | |
| | e35:0 0 0 + - | | |
| | e37:0 0 0 - + | | |
| | f25: 0 0 0 | | |
| TF^+ | e1: 0 - + · · | | |
| | e16:0 0 + \cdot \cdot | | |
| | $e_{35:0} 0 + - \cdot$ | | |
| | $e_{37:0} 0 - + \cdot$ | | |
| | t25: 0 0 | | |

| Actual | e1: - · · e16: 0 0 e35: 0 0 |
|--|--|
| | e35:0 0 |
| | 633.0 0 |
| | e37· - 0 |
| | f25: - 0 |
| GY - | e1: 0 - · |
| | e16:0 - · |
| | e35:0 - + |
| | e37:-+- |
| - | f25: 0 · · · |
| I_{ep}^+ | e1: 0 – · |
| | e16:0 – · |
| | $e_{35:0} - +$ |
| | e37:-+- |
| n+ | 125: 0 |
| κ_{ep} | el: 0 0 - |
| | $e_{16:0} 0 - $ |
| | $e_{35}: 0 \ 0 -$ |
| | $e_{27:0} = +$ |
| | 140.00 |
| R^{-} | e1· 0 |
| R_{brine}^{-} | e1: $0 - +$ |
| R^{-}_{brine} | e1: $0 - +$ e16: $0 - +$ e35: $0 +$ |
| R_{brine}^{-} | e1: $0 - +$ e16: $0 - +$ e35: $0 0 +$ e37: $0 0 -$ |
| R_{brine}^{-} | e1: $0 - +$ e16: $0 - +$ e35: $0 0 +$ e37: $0 0 -$ f25: $0 -$ $t_f = 1960(s)$ |
| R_{brine}^{-} Step 3 (t – Actual | e1: $0 - +$ e16: $0 - +$ e35: $0 0 +$ e37: $0 0 -$ f25: $0 - \cdot$ e1f = 1960(s) e1: $- 0$ e16: $0 0$ |
| R_{brine}^{-} | $\begin{array}{c} \text{e1: } 0 - + \\ \text{e16: } 0 - + \\ \text{e35: } 0 \ 0 + \\ \text{e37: } 0 \ 0 - \\ \text{f25: } 0 - \\ \text{f25: } 0 - \\ \text{f25: } 0 - \\ \text{e37: } 0 \\ \text{e16: } 0 \\ \text{e35: } - \\ \text{e37: } 0 \end{array}$ |
| R_{brine}^{-} | $\begin{array}{c} {\rm e1:} \ 0 - + \\ {\rm e16:} \ 0 - + \\ {\rm e35:} \ 0 \ 0 + \\ {\rm e37:} \ 0 \ 0 - \\ {\rm f25:} \ 0 - \\ {\rm f25:} \ 0 - \\ {\rm e16:} \ 0 \\ {\rm e16:} \ 0 \\ {\rm e35:} - \\ {\rm e37:} - \ 0 \\ {\rm f25:} \ 0 \end{array}$ |
| R_{brine}^{-} Step 3 (t – Actual | $\begin{array}{c} \text{e1: } 0 - + \\ \text{e16: } 0 - + \\ \text{e35: } 0 \ 0 + \\ \text{e37: } 0 \ 0 - \\ \text{f25: } 0 - \\ \text{f25: } 0 - \\ \text{f16: } 0 \\ \text{e16: } 0 \\ \text{e16: } 0 \\ \text{e16: } 0 \\ \text{e35: } - \\ \text{e37: } - \\ \text{e37: } - \\ \text{f25: } 0 \\ \text{e1: } 0 $ |
| $\frac{R_{brine}}{R_{brine}}$ Step 3 (t – Actual | $\begin{array}{c} {\rm e1:} \ 0 - + \\ {\rm e16:} \ 0 - + \\ {\rm e35:} \ 0 \ 0 + \\ {\rm e37:} \ 0 \ 0 - \\ {\rm f25:} \ 0 - \\ {\rm f25:} \ 0 - \\ {\rm e16:} \ 0 \\ {\rm e35:} - \\ {\rm e37:} - \\ {\rm e37:} - \\ {\rm e16:} \ 0 - \\ { $ |
| R_{brine}^{-} Step 3 ($t - Actual$ | $\begin{array}{c} \mathrm{e1:} \ 0 \ - \ + \\ \mathrm{e16:} \ 0 \ - \ + \\ \mathrm{e16:} \ 0 \ - \ + \\ \mathrm{e35:} \ 0 \ 0 \ - \\ \mathrm{e37:} \ 0 \ - \\ \mathrm{e15:} \ - \ 0 \\ \mathrm{e16:} \ 0 \ - \ \cdot \\ \mathrm{e16:} \ 0 \ - \ \cdot \\ \mathrm{e16:} \ 0 \ - \ \cdot \\ \mathrm{e16:} \ 0 \ - \ - \ + \\ \mathrm{e16:} \ 0 \ - \ - \ + \\ \mathrm{e16:} \ 0 \ - \ - \ + \\ \mathrm{e16:} \ 0 \ - \ - \ + \\ \mathrm{e16:} \ 0 \ - \ - \ + \\ \mathrm{e16:} \ 0 \ - \ - \ + \\ \mathrm{e16:} \ 0 \ - \ - \ + \\ \mathrm{e16:} \ 0 \ - \ - \ + \\ \mathrm{e16:} \ 0 \ - \ - \\ \mathrm{e16:} \ 0 \ - \ - \ + \\ \mathrm{e16:} \ 0 \ - \ - \ + \\ \mathrm{e16:} \ 0 \ - \ - \ + \\ \mathrm{e16:} \ 0 \ - \ - \ + \\ \mathrm{e16:} \ 0 \ - \ - \ + \\ \mathrm{e16:} \ 0 \ - \ - \ + \\ \mathrm{e16:} \ 0 \ - \ - \ + \\ \mathrm{e16:} \ 0 \ - \ - \ + \\ \mathrm{e16:} \ - \ - \ - \ + \\ \mathrm{e16:} \ - \ - \ - \ + \ - \\ \mathrm{e16:} \ - \ - \ - \ - \ - \ - \ - \ - \ - \ $ |
| $\frac{R_{brine}^{-}}{\text{Actual}}$ | $\begin{array}{c} \mathrm{e1:} \ 0 \ - \ + \\ \mathrm{e16:} \ 0 \ - \ + \\ \mathrm{e35:} \ 0 \ 0 \ + \\ \mathrm{e37:} \ 0 \ 0 \ - \\ \mathrm{e37:} \ 0 \ - \\ \mathrm{e16:} \ 0 \ - \\ \mathrm{e16:} \ 0 \ - \\ \mathrm{e35:} \ - \\ \mathrm{e35:} \ - \\ \mathrm{e35:} \ - \\ \mathrm{e35:} \ - \\ \mathrm{e37:} $ |
| R_{brine}^{-} Step 3 (t – Actual | $\begin{array}{c} \mathrm{e1:} \ 0 \ - \ + \\ \mathrm{e16:} \ 0 \ - \ + \\ \mathrm{e35:} \ 0 \ 0 \ + \\ \mathrm{e37:} \ 0 \ 0 \ - \\ \mathrm{f25:} \ 0 \ - \\ \mathrm{f25:} \ 0 \ - \\ \mathrm{e37:} \ - \ 0 \\ \mathrm{e16:} \ 0 \ - \\ \mathrm{e37:} \ - \\ \mathrm{e16:} \ 0 \ - \\ \mathrm{e35:} \ - \\ \mathrm{e37:} \ - \\ $ |
| $\frac{R_{brine}^{-}}{\text{Step 3}(t - Actual}$ $\frac{GY^{-}}{I_{ep}^{+}}$ | $\begin{array}{c} \mathrm{el:} \ 0 \ - \ + \\ \mathrm{el6:} \ 0 \ - \ + \\ \mathrm{el6:} \ 0 \ - \ + \\ \mathrm{el7:} \ 0 \ 0 \ - \\ \mathrm{el7:} $ |
| $\frac{R_{brine}^{-}}{GY^{-}}$ | $\begin{array}{c} \mathrm{e1:} \ 0 \ - \ + \\ \mathrm{e16:} \ 0 \ - \ + \\ \mathrm{e16:} \ 0 \ - \ + \\ \mathrm{e35:} \ 0 \ 0 \ + \\ \mathrm{e37:} \ 0 \ - \\ \mathrm{e125:} \ 0 \ - \\ \mathrm{e16:} \ 0 \ 0 \\ \mathrm{e35:} \ - \\ \mathrm{e37:} \ - \\ \mathrm{e16:} \ 0 \ - \\ \mathrm{e35:} \ 0 \ - \\ \mathrm{e15:} \ 0 \ - \\ \mathrm{e16:} \ 0 \ - \\ \mathrm{e35:} \ - \\ \mathrm{e16:} \ 0 \ - \ 0$ |
| $\frac{R_{brine}^{-}}{GY^{-}}$ | $\begin{array}{c} \mathrm{e1:} \ 0 - + \\ \mathrm{e16:} \ 0 - + \\ \mathrm{e16:} \ 0 - + \\ \mathrm{e35:} \ 0 \ 0 + \\ \mathrm{e37:} \ 0 \ 0 - \\ \mathrm{e12:} \ 0 - \cdot \\ \mathrm{e16:} \ 0 \\ \mathrm{e35:} - \cdot \\ \mathrm{e37:} - \ 0 \\ \mathrm{e15:} \ 0 \\ \mathrm{e15:} - \cdot \\ \mathrm{e16:} \ 0 - \cdot \\ \mathrm{e35:} - + \\ \mathrm{e37:} - \ 0 \\ \mathrm{e15:} \ 0 - \cdot \\ \mathrm{e16:} \ 0 - \cdot \\ \mathrm{e35:} - + \\ \mathrm{e37:} - + - \\ \mathrm{e16:} \ 0 - \cdot \\ \mathrm{e16:} \ 0 - \cdot \\ \mathrm{e15:} \ 0 - \cdot \\ \mathrm{e16:} \ 0 - \cdot \\ \mathrm{e15:} \ 0 - \mathrm{e15:} \ 0 - \\ \mathrm{e15:} \ 0 - \mathrm{e15:} \ $ |
| $\frac{R_{brine}^{-}}{\text{Step 3}(t - Actual}$ $\frac{GY^{-}}{I_{ep}^{+}}$ | $\begin{array}{c} \mathrm{e1:} \ 0 - + \\ \mathrm{e16:} \ 0 - + \\ \mathrm{e35:} \ 0 \ 0 + \\ \mathrm{e37:} \ 0 \ 0 - \\ \mathrm{e37:} \ 0 \ - \\ \mathrm{e125:} \ 0 - \\ \mathrm{e16:} \ 0 \\ \mathrm{e35:} - \\ \mathrm{e16:} - \\ \mathrm{e35:} - \\ \mathrm{e15:} - \\ \mathrm{e15:}$ |
| $\frac{R_{brine}^{-}}{Actual}$ $\frac{GY^{-}}{I_{ep}^{+}}$ $p + \frac{GY^{-}}{Bc}$ | $\begin{array}{c} \text{e1: } 0 - + \\ \text{e16: } 0 - + \\ \text{e35: } 0 + \\ \text{e37: } 0 - \\ \text{f25: } 0 - \\ \text{f25: } 0 - \\ \text{f25: } 0 - \\ \text{e16: } 0 \\ \text{e35: } - \\ \text{e37: } - \\ - \\ \text{e37: } - \\ - \\ - \\ - \\ - \\ - \\ - \\ - \\ - \\ -$ |
| $\frac{R_{brine}^{-}}{Actual}$ $\frac{GY^{-}}{I_{ep}^{+}}$ R_{ep}^{+} | $\begin{array}{c} \mathrm{el:} \ 0 \ - \ + \\ \mathrm{el6:} \ 0 \ - \ + \\ \mathrm{el6:} \ 0 \ - \ + \\ \mathrm{el5:} \ 0 \ 0 \ - \\ \mathrm{el5:} \ 0 \ - \ $ |
| $\frac{R_{brine}^{-}}{Actual}$ $\frac{GY^{-}}{I_{ep}^{+}}$ $\frac{R_{ep}^{+}}{R_{ep}^{+}}$ | $\begin{array}{c} \mathrm{e1:} \ 0 \ - \ + \\ \mathrm{e16:} \ 0 \ - \ + \\ \mathrm{e16:} \ 0 \ - \ + \\ \mathrm{e35:} \ 0 \ 0 \ - \\ \mathrm{e37:} \ 0 \ - \\ \mathrm{e16:} \ 0 \ - \\ \mathrm{e35:} \ - \\ \mathrm{e16:} \ 0 \ - \\ \mathrm{e35:} \ - \\ \mathrm{e35:} \ - \\ \mathrm{e35:} \ - \\ \mathrm{e16:} \ 0 \ - \\ \mathrm{e35:} \ - \\ \mathrm{e15:} \ 0 \ - \\ \mathrm{e16:} \ 0 \ - \\ \mathrm{e35:} \ - \\ \mathrm{e15:} \ 0 \ 0 \ - $ |
| $\frac{R_{brine}^{-}}{Actual}$ $\frac{GY^{-}}{I_{ep}^{+}}$ $\frac{R_{ep}^{+}}{R_{ep}^{+}}$ | $\begin{array}{c} \mathrm{el:} \ 0 \ - \ + \\ \mathrm{el6:} \ 0 \ - \ + \\ \mathrm{el6:} \ 0 \ - \ + \\ \mathrm{el6:} \ 0 \ - \ + \\ \mathrm{el7:} \ 0 \ - \ 0 $ |

Fig. 9. Qualitative fault isolation for fault GY^- with fault size 5% reduces the possible set of faults to three candidates.

e1: 0 0 e16:0 0 e35:0 0 e37:- 0 f25: e1: 0 e16:0 - · e35:0 - + e37:- + - · f25: 0 · · · e1: 0 - · e16:0 - · · $e35:0 - + \cdot$ e37:- + - · f25: 0 · · · e1: 0 0 – \cdot e16:0 0 - · e35:0 0 - + e37:0 - + f25: 0 0 \cdot \cdot e1: 0 - + $e16:0 - + \cdot$ e35:0 0 + e37:0 0 - + f25: 0 e1: $0 - + \cdot$ e16:0 0 + \cdot e35:0 0 + e37:0 0 - + f25:00 · · e1: 0 0 - + e16:0 0 0 + e35:0 0 0 + e37:0 0 0 - + f25:000 · e1: 0 - +e16:0 0 + \cdot e35:0 0 + e37:0 0 - + f25:00 ·

of the parameter estimation scheme for this fault are given in Table II. The parameter estimation provides the least error for the GY fault hypothesis, therefore, the diagnosis system correctly returns GY^{-} , i.e., a decrease in pump efficiency as the fault candidate, with a multiplicative factor of 0.934, which is close to the actual fault magnitude of -5%.

B. Comprehensive fault isolation results for the RO system

Table II presents the comprehensive diagnosis results for selected faults in the RO system. As with the detailed example discussed earlier, the faults are introduced in the primary mode of operation although the actual time of fault occurrence varies. The fault magnitudes are chosen to ensure detection (after some delay) and the generation of the correct deviated signals using the signal-to-symbol transformation scheme used. The sensitivity of the residual to a change in parameter value is parameter dependent but this paper does not explore this sensitivity, and the related statistical detection performance issues. A discussion of detection and estimation statistics was presented in [10].

For each scenario, the qualitative fault isolation scheme reduces the initial candidate set considerably, and parameter estimation converges to the correct fault candidate. The estimated parameter values are also quite acceptable for all scenarios. This demonstrates the effectiveness of the diagnosis methodology in isolating and identifying various faults in the RO system. In future, we will conduct more systematic experiments on fault sensitivity and robustness to measurement noise and modeling errors.

| Fault | $t - t_f$ | Step | Symbolic | Candidate set + parameter estimation |
|--|-----------|------|---------------------------|---|
| $R^+_{memb}, 5\%$ $t_f: 20000$ | 800 | 0 | $F_{perm}(f25):(-,\cdot)$ | $C_{c}^{+}, C_{memb}^{+}, I_{fp}^{+}, I_{ep}^{-}, R_{brine}^{-}, TF^{+}, R_{pipe}^{-}, R_{memb}^{+}, R_{fp}^{+}, R_{ep}^{-}, GY^{-}$ |
| | 7200 | 1 | $P_{back}(e1):(+,\cdot)$ | $I_{fp}^{-}, TF^{+}, R_{pipe}^{-}, R_{memb}^{+}, R_{fp}^{+}, R_{ep}^{-}$ |
| | 8280 | 2 | $P_{memb}(e16):(+,\cdot)$ | $R_{pipe}^-, R_{memb}^+, R_{ep}^-$ |
| | | | | parameter estimation selects R^+_{memb} , indicates change by 1.042 |
| $GY^-, 5\%$ $t_f: 17500$ | 200 | 0 | $P_{pump}(e37):(-,\cdot)$ | $C_{c}^{+}, C_{memb}^{+}, I_{fp}^{+}, I_{ep}^{+}, R_{brine}^{-}, TF^{+}, R_{pipe}^{-}, R_{memb}^{-}, C_{k}^{+}, R_{fp}^{+}, R_{ep}^{+}, GY^{-}$ |
| | 880 | 1 | $F_{perm}(f25):(-,\cdot)$ | $I_{fp}^+, I_{ep}^+, R_{brine}^-, TF^+, R_{fp}^+, R_{ep}^+, GY^-$ |
| | 1240 | 2 | $P_{back}(e1):(-,\cdot)$ | $I_{ep}^+, R_{brine}^-, R_{ep}^+, GY^-$ |
| | 1960 | 3 | $K(e35)$: $(-, \cdot)$ | I_{ep}^+, R_{ep}^+, GY^- |
| | | | | parameter estimation: GY^- changed by 0.934 |
| $R_{ep}^+, 35\%$ t $\epsilon : 20000$ | 88 | 0 | $P_{pump}(e37):(-,\cdot)$ | $C_{c}^{+}, C_{memb}^{+} I_{fp}^{+}, I_{ep}^{+}, R_{brine}^{-}, TF^{+}, R_{pipe}^{-}, R_{memb}^{-}, C_{k}^{+}, R_{fp}^{+}, R_{ep}^{+}, GY^{-}$ |
| | 640 | 1 | $F_{perm}(f25):(-,\cdot)$ | $I_{fp}^{+}, I_{ep}^{+}, R_{brine}^{-}, TF^{+}, R_{fp}^{+}, R_{ep}^{+}, GY^{-}$ |
| | 720 | 2 | $P_{back}(e1):(-,\cdot)$ | $I_{ep}^+, R_{brine}^-, R_{ep}^+, GY^-$ |
| , | 960 | 3 | $P_{pump}(e37):(-,-)$ | $R_{brine}^{-}, R_{ep}^{+}$ |
| | 4640 | 4 | $K(e35)$: $(-, \cdot)$ | R_{ep}^+ |
| | | | | parameter estimation: R_{ep}^+ changed by 0.374 |
| $R^+_{pipe}, 15\%$ $t_f: 18000$ | 640 | 0 | $P_{back}(e1):(-,\cdot)$ | $C_{c}^{-}, C_{memb}^{-}, I_{fp}^{+}, R_{brine}^{+}, TF^{+}, R_{pipe}^{+}, R_{memb}^{+}, C_{k}^{+}, R_{fp}^{-}, R_{ep}^{+}, GY^{+}$ |
| | 800 | 1 | $P_{memb}(e16):(-,\cdot)$ | $R_{brine}^-, TF^+, R_{pipe}^+, R_{ep}^+,$ |
| | | | | parameter estimation: R_{pipe}^+ changed by 1.134 |
| $C_{memb}^{-}, 10\%$ $t_f: 19600$ | 360 | 0 | $F_{perm}(f25):(-,\cdot)$ | $C_{c}^{-}, C_{memb}^{-}, I_{fp}^{+}, R_{brine}^{+}, TF^{+}, R_{pipe}^{+}, R_{memb}^{-}, C_{k}^{+}, R_{fp}^{-}, R_{ep}^{+}, GY^{+}$ |
| | 480 | 1 | $P_{back}(e1):(-,\cdot)$ | $C_{memb}^- R_{brine}^+, TF^+, GY^+$ |
| | 8680 | 2 | $P_{memb}(e16):(-,\cdot)$ | $C^{-}_{memb} R^{+}_{brine}, GY^{+}$ |
| | | | | parameter estimation: C_{memb}^{-} changed by 0.856 |

TABLE II

COMPREHENSIVE DIAGNOSIS RESULTS FOR SELECTED FAULTS IN THE RO SYSTEM.

V. DISCUSSION AND CONCLUSIONS

The results presented above demonstrate the effectiveness of our model-based fault diagnosis scheme. The HBG modeling paradigm provides a topological representation that captures the continuous system dynamics along with discrete mode changes that correspond to physical reconfiguration of the system. It also allows for modeling multiple physical paradigms using a compact representation. The innovation in our modeling work is the ability to capture the change in conductivity as the water circulates in the RO system, and also seamlessly model the changes that occur as the mode of system operation (cyclic behavior from primary to secondary to purge mode) changes. The model was validated with real data that was obtained from the JSC testbed. Another significant advantage of the HBG paradigm is that it captures the temporal and causal relations between system variables and component parameters, allowing for very efficient qualitative models for diagnosis of dynamic systems. This helps overcome some of the limitations that have been observed for analysis of multiplicative faults in systems with complex nonlinear behaviors. As the experimental results in the previous section demonstrate, the qualitative scheme is ambiguous, but once the fault set is reduced to a small size, quantitative estimation techniques can be applied to uniquely isolate the fault and compute the magnitude of change.

For quick detection of fault transients with small magnitudes, and reliable symbol generation to enable precise isolation, one has to carefully tune the statistical parameters of the fault detectors and symbol generators. A statistically significant evaluation of the detection and isolation performance with respect to different fault sizes will be presented in future work.

Our diagnosis scheme can be applied online, and the time difference between actual fault occurrence, fault detection, isolation, and estimation are demonstrated in our experimental results. Estimation of the fault magnitude is critical for the model-predictive fault-adaptive control techniques we have been developing in other work [6], [23]. Another future task will be to study the effect of delays in the fault-adaptive control task.

In conclusion, we are developing new online technology for embedded hybrid systems that can be employed for fault diagnosis and fault-adaptive control of complex systems. Further development of this technology will provide the proof of concept that advanced control techniques can form the backbone for autonomy in future long-duration missions that NASA deploys.

ACKNOWLEDGMENTS

This work was supported in part through grants from the NASA-IS program (Contract number: NAS2-37143) and the NASA-ALS program (Contract number: NCC 9-159). We acknowledge the help provided by the members of the ALS group at NASA Johnson Space Center (JSC), primarily, Pete Bonasso, David Kortenkamp, Wen-Ching Lee, and other members of the Lockheed group at NASA JSC.

BIOGRAPHIES

Gautam Biswas is Professor of Computer Science and Engineering, and Management of Technology at Vanderbilt University and a senior research scientist at the Institute for Software Integrated Systems (ISIS) at Vanderbilt University. He has a Ph.D. degree in Computer Science from Michigan State University in E. Lansing, MI.

Prof. Biswas conducts research in Intelligent Systems with primary interests in hybrid modeling, simulation, and analysis of complex embedded systems, and their applications to diagnosis and fault-adaptive control. As part of this work, he is working on fault-adaptive control of fuel transfer systems for aircraft and UCAVs, and the water recovery system of the Bio-Plex system. He is also initiating new projects in distributed monitoring and condition-based maintenance. In other research projects, he is also involved in developing simulation-based environments for learning and instruction. His research is currently supported by funding from NASA, DARPA, and the NSF.

Dr. Biswas is an associate editor of the IEEE Transactions on Systems, Man, and Cybernetics, the IEEE Transactions on Knowledge and Data Engineering, and the Journal of Applied Intelligence. He has served on the Program Committee of a number of conferences. He is a senior member of the IEEE Computer Society, ACM, AAAI, and the Sigma Xi Research Society.

- *Eric-Jan Manders* is a research associate in the Institute of Software Integrated Systems (ISIS) and the Department of Electrical Engineering and Computer Science at Vanderbilt University. He received his 'Ingenieur' (Ir.) degree, in Electrical Engineering from Delft University of Technology, The Netherlands, in 1991, and a Ph.D. in Electrical Engineering from Vanderbilt university in 2003. His research interests include knowledge-based signal processing, model-based fault detection and isolation, instrumentation, and distributed real-time embedded systems.
- *John Ramirez* is a member of the research staff at the Institute of Software Integrated Systems (ISIS) He obtained a B.S. in Electrical Engineering from Tennessee Technological University in 2001, and an M.S. in Computer Science from Vanderbilt University in 2003.
- *Nagabhusan Mahadevan* is a member of the research staff at the Institute of Software Integrated Systems (ISIS) at Vanderbilt University. He received a M.S. in Chemical Engineering from the University of South Carolina, Columbia.
- *Sherif Abdelwahed* is a research assistant professor at the Institute of Software Integrated Systems (ISIS) at Vanderbilt University. He received his Ph.D. in Electrical and Computer Engineering from the University of Toronto in 2002. His research interests are in the areas of discrete event and hybrid systems, design and verification of supervisory controllers, and analysis of large-scale systems.

REFERENCES

- B.E. Duffield and A.J. Hanford, "Advanced life support requirements document," Tech. Rep. JSC-38571, Rev. B, NASA-Lyndon B. Johnson Space Center, Houston, TX, Sept. 2002.
- [2] National Aeronautics and Space Administration, "The vision for space exploration," NP-2004-01-334-HQ, Feb. 2004.
- [3] R. Isermann, "Supervision, fault-detection and fault-diagnosis methods an introduction," *Control Engineering Practice*, vol. 5, no. 5, pp. 639–652, May 1997.
- [4] P. Bonasso, D. Kortenkamp, and C. Thronesbery, "Intelligent control of a water recovery system: Three years in the trenches," *AI Magazine*, pp. 19–43, 2003.
- [5] P. J. Mosterman and G. Biswas, "A theory of discontinuities in dynamic physical systems," *J Franklin Institute*, vol. 334B, no. 6, pp. 401–439, 1997.
- [6] G. Karsai, G. Biswas, T. Pasternak, S. Narasimhan, G. Peceli, G. Simon, and T. Kovacshazy, "Towards fault-adaptive control of complex dynamical systems," in *Software-Enabled Control – Information Technology for Dynamical Systems*, T. Samad and G. Balas, Eds., chapter 17, pp. 347–368. Wiley-IEEE press, NJ, 2003.
- [7] J. J. Gertler, *Fault Detection and Diagnosis in Engineering Systems*, Marcel Dekker, Inc., New York, NY, 1998.
- [8] D. C. Karnopp, D. L. Margolis, and R. C. Rosenberg, *Systems Dynamics: Modeling and Simulation of Mechatronic Systems*, John Wiley & Sons, Inc., New York, third edition, 2000.
- [9] J. Chen and R. J. Patton, *Robust Model-Based Fault Diagnosis for Dynamic Systems*, Kluwer Academic publishers, Boston, MA USA, 1998.
- [10] E.-J. Manders and G. Biswas, "FDI of abrupt faults with combined statistical detection and estimation and qualitative fault isolation," in *Proc. 5th IFAC Symp Fault Detection Supervision Safety Technical Processes*, Washington, DC, June 2003, pp. 347–352.
- [11] P. J. Mosterman and G. Biswas, "Diagnosis of continuous valued systems in transient operating regions," *IEEE Trans. Syst., Man Cybern. A*, vol. 29, no. 6, pp. 554–565, 1999.
- [12] D. Weld and J. de Kleer, Eds., *Readings in Qualitative Reasoning about Physical Systems*, Morgan Kaufmann, San Mateo, CA, 1990.
- [13] Gautam Biswas and Xudong Yu, "A formal modeling scheme for continuous systems: Focus on diagnosis," in *Proc. Thirteenth Int. Joint Conf. Artificial Intelligence*, Chambery, France, Aug. 1993, pp. 1474–1479.
- [14] M. Basseville and I. V. Nikiforov, *Detection of abrupt changes: theory and applications*, Prentice Hall, Englewood Cliffs, New Jersey, 1993.
- [15] F. Gustafsson, *Adaptive filtering and change detection*, John Wiley & Sons, Ltd, United Kingdom, 2000.
- [16] K. Brammer and G. Siffling, Kalman-Bucy Filters, Artec House, Norwood MA, 1989.
- [17] G. Biswas, G. Simon, N. Mahadevan, S. Narasimhan, J. Ramirez, and G. Karsai, "A robust method for hybrid diagnosis of complex systems," in *Proc. 5th IFAC Symp Fault Detection Supervision Safety Technical Processes*, Washington, DC, June 2003, pp. 1125–1131.
- [18] E.-J. Manders, S. Narasimhan, G. Biswas, and P. J. Mosterman, "A combined qualitative/quantitative approach for fault isolation in continuous dynamic systems," in *Proc. 4th IFAC Symp Fault Detection Supervision Safety Technical Processes*, Budapest, Hungary, June 2000, pp. 1074–1079.
- [19] D. Kortenkamp and S. Bell, "BioSim: An integrated simulation of an advanced life support system for intelligent control research," in *Proc. of the 7th Intl. Symp. on Artificial Intelligence, Robotics and Automation in Space*, 2003.
- [20] K. D. Pickering, K. R. Wines, G. M. Pariani, L. A. Franks, J. Yeh, B. W. Finger, M. L. Campbell,

C. E. Verostko, C. Carrier, J. C. Gandhi, and L. M. Vega, "Early results of an integrated water recovery system test," in *Proc 29th Int Conf Environmental Sys*, 2001.

- [21] P. J. Feenstra, P. J. Mosterman, G. Biswas, and P. C. Breedvald, "Bond graph modeling procedures for fault detection and isolation of complex flow processes," in *Proc. 2001 Int Conf Bond Graph Modeling Simulation*, Jan. 2000.
- [22] G. Karsai, J. Sztipanovits, A. Ledeczi, and T. Bapty, "Model-integrated development of embedded software," *Proc IEEE*, vol. 91, no. 1, pp. 145–164, Jan. 2003.
- [23] S. Abdelwahed, J. Wu, G. Biswas, J. Ramirez, and E.-J. Manders, "On-line hierarchical fault adaptive control for advanced life support systems," in *Proc 32nd Int Conf Environmental Sys*, 2004, To Appear.