

Comprehensive Diagnosis of Continuous Systems Using Dynamic Bayes Nets

Indranil Roychoudhury, Gautam Biswas, and Xenofon Koutsoukos

Institute for Software Integrated Systems (ISIS)

Department of Electrical Engineering and Computer Science

Vanderbilt University, Nashville, TN, 37235, USA

Email: {indranil.roychoudhury, gautam.biswas, xenofon.koutsoukos}@vanderbilt.edu

Abstract

Fault diagnosis is essential for guaranteeing safe and reliable operation of complex engineering systems. Our work focuses on diagnosis of parametric faults in the components of dynamic systems, whose temporal profile can be categorized as incipient (slow) or abrupt (fast). The diagnosis of abrupt and incipient faults using qualitative approaches is challenging, since in many situations, these faults produce similar qualitative effects. Quantitative estimation methods may provide more discriminatory power, but these approaches can be computationally infeasible for large systems with nonlinearities and complex dynamics. In this paper, we combine a qualitative fault isolation scheme with an Dynamic Bayes net-based particle filtering approach for the comprehensive diagnosis of incipient and abrupt faults in continuous systems. We also present experimental results to demonstrate the effectiveness of our approach when applied to a two-tank system.

1 Introduction

Detection, isolation, and identification of faults in system components is essential for guaranteeing safe, reliable, and efficient operation of complex engineering systems. Some of these faults are attributed to degradations, and modeled as *incipient* faults, i.e., slow drifts in system parameter values over time. Other faults manifest as quick changes in component parameter values and are modeled as *abrupt* faults, i.e., changes in parameter values that are fast in comparison to the system dynamics, and approximated as a step change.

In the past, we have successfully diagnosed abrupt faults using qualitative schemes [Mosterman and Biswas, 1999]. However, qualitative diagnosis schemes for both incipient and abrupt faults may suffer from the ambiguity problem, i.e., the inability to discriminate among fault hypotheses. Quantitative approaches produce more precise diagnoses, but, for large systems with complex dynamics, these quantitative approaches can be computationally expensive. In this paper, we extend our earlier work [Roychoudhury *et al.*, 2006] to the comprehensive diagnosis of *both* incipient and abrupt faults in continuous dynamic systems through the integration of our

qualitative diagnosis scheme [Mosterman and Biswas, 1999] with an Dynamic Bayes net-based *particle filtering* approach.

Dynamic Bayes Nets (DBNs) exploit the conditional independence among variables to provide a compact and factored representation of a dynamic system and allow arbitrary uncertainty models of the dynamic process and measurement [Murphy, 2002]. Hence, DBN-based tracking approaches need not conform to the restrictive assumption of normal distributions for noise and modeling errors. DBN schemes have been developed for several fault diagnosis problems [Lerner *et al.*, 2000; Murphy, 2002]. This, however, makes the calculation of posterior probabilities computationally expensive, as, in many cases, no analytic closed form solutions for these probabilities exist. Particle filters (PFs) are a state-of-the-art inference mechanism using DBNs [Koller and Lerner, 2001] that help overcome the inability to derive analytic solutions for posterior probabilities. Moreover, for large systems, fairly significant gains in computational efficiency can be achieved by leveraging the sparseness and compactness of the DBN structures over traditional state-space based methods. Especially, in DBN models of faulty systems, the introduction of fault parameters typically results in very few additional links.

In our approach, we use PFs applied to the DBN model of the nominal system to generate estimates of nominal system behavior that are robust to measurement noise and modeling error. A statistically significant non-zero residual value implies a fault, and the qualitative fault isolation (Qual-FI) scheme generates and prunes fault hypotheses as measurement deviations are observed. Then, the quantitative fault isolation and identification (Quant-FII) scheme is invoked to further refine the fault hypotheses. A faulty DBN is generated for each remaining fault candidate by including the fault parameter as a stochastic variable in the DBN, and a separate PF scheme is run on each faulty DBN to track the faulty system behavior. As the Qual-FI scheme continues to refine its fault hypotheses, the PFs tracking the measurements using the inconsistent fault models are terminated. Also, if the measurements estimated by the PF applied to a particular fault model significantly deviates from the observed faulty measurements, that fault candidate is deemed inconsistent and removed from the set of possible faults. Eventually, the PFs using the true fault model converges to the observed faulty measurements, and estimates the value of the fault parameter. This efficient pruning of inconsistent fault hypotheses based

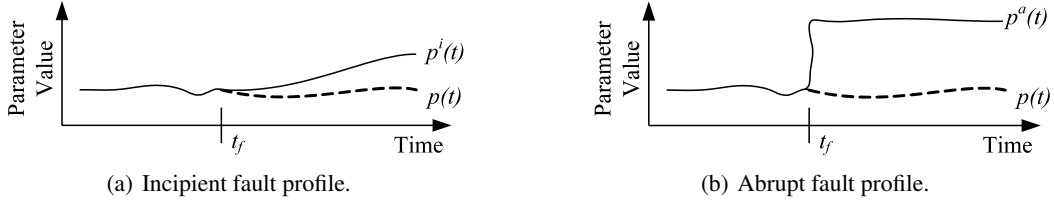


Figure 1: Fault profiles.

on the qualitative and quantitative analysis of measurements helps in fast diagnosis of the true fault.

The paper is organized as follows. Section 2 presents mathematical models of incipient and abrupt faults. Section 3 describes our diagnosis architecture. The different models used in our approach are described in Section 4. Section 5 discusses our fault detection approach, while Section 6 explains the fault isolation and identification scheme for incipient and abrupt faults in detail. Section 7 presents experimental results, and conclusions are presented in Section 8.

2 Incipient and Abrupt Faults

The mathematical models for incipient and abrupt faults used in our approach are defined below. In general, incipient and abrupt faults can be additive and multiplicative. However, in this paper, we focus on parametric multiplicative faults, which are hard to analyze because they directly affect the system dynamics.

2.1 Incipient Faults

An incipient fault is a slow change in a system parameter. Hence, we model this fault as a linear, additive, drift term, $d(t)$, added to the nominal component parameter value function, $p(t)$. Since incipient faults are slow changes, we approximate $d(t)$ as a linear function with a constant slope. Fig. 1(a) shows an incipient fault profile.

Definition 1 (Incipient fault) *An incipient fault profile is characterized by a gradual, slow drift in the corresponding component parameter value. The temporal profile of a parameter with an incipient fault, $p^i(t)$, is given by:*

$$p^i(t) = \begin{cases} p(t) & t \leq t_f \\ p(t) + d(t) = p(t) + \sigma_p^i \times (t - t_f) & t > t_f, \end{cases} \quad (1)$$

where $p(t)$ represents the nominal parameter value, $d(t)$ is a drift function, and t_f is the time point of fault occurrence.

2.2 Abrupt Faults

An abrupt fault is modeled as an addition of a bias term, $b(t)$, to the nominal parameter value, $p(t)$. Typically, abrupt faults are very fast changes, and so, the bias term is modeled as an additive step function (see Fig. 1(b)). We assume the magnitude of this bias term to be constant.

Definition 2 (Abrupt fault) *An abrupt fault profile is characterized by a fast change in the component parameter value.*

The temporal profile of a parameter with an abrupt fault, $p^a(t)$, is given by:

$$p^a(t) = \begin{cases} p(t) & t < t_f \\ p(t) + b(t) = p(t) + \sigma_p^a & t \geq t_f, \end{cases} \quad (2)$$

where $p(t)$ denotes the nominal parameter value, $b(t)$ is a bias term, and t_f is the time point of fault occurrence.

3 Diagnosis Architecture

Our combined model-based approach for diagnosing abrupt and incipient faults, like traditional model-based diagnosis schemes [Gertler, 1998], has three primary components: (i) fault detection, (ii) fault isolation, and (iii) fault identification, as summarized below. The architecture of our diagnosis methodology is shown in Fig. 2.

Fault Detection: The dynamic nominal behavior of the system is tracked by a PF-based observer scheme [Koller and Lerner, 2001] based on a DBN model of the nominal system, i.e., the state variables and measurements made on the system are modeled as stochastic variables but the system parameters are considered to be deterministic and defined by their nominal functions. Like other observer schemes, the PF generates estimates of the state variables, $\hat{x}(t)$, and measurements, $\hat{y}(t)$. The fault detector monitors each measurement residual, $r(t) = y(t) - \hat{y}(t)$, at each time step, where $y(t)$ is a measured variable at time t , and $\hat{y}(t)$ is the value of the measurement estimated by the PF. Ideally, $r(t) \neq 0$ should imply a fault and trigger the fault isolation scheme, but to accommodate measurement noise and modeling errors we employ a statistical testing scheme that balances detection sensitivity against false alarms, and a fault is detected if a non-zero residual is statistically significant.

Fault Isolation: Once a fault is detected, the fault isolation module is activated. Fault isolation is performed by running a Qual-FI scheme that uses the symbolic values of measurement deviations along with a Quant-FII approach that is based on a DBN-based PF scheme. The Qual-FI scheme for abrupt and incipient faults is described in Section 6.1. Once the number of fault hypotheses is less than a pre-defined threshold, k , or the Qual-FI scheme has run for s steps, where s is also a design parameter, we invoke the Quant-FII scheme.

The Quant-FII scheme starts with a separate DBN model for each fault candidate listed in the qualitative fault hypothesis set. The extended DBN includes the fault parameter as a stochastic variable in the DBN, and the corresponding PF tracks the faulty system measurements from the time of fault detection. If the system is diagnosable, the measurements estimated by the PF using the true fault model will converge

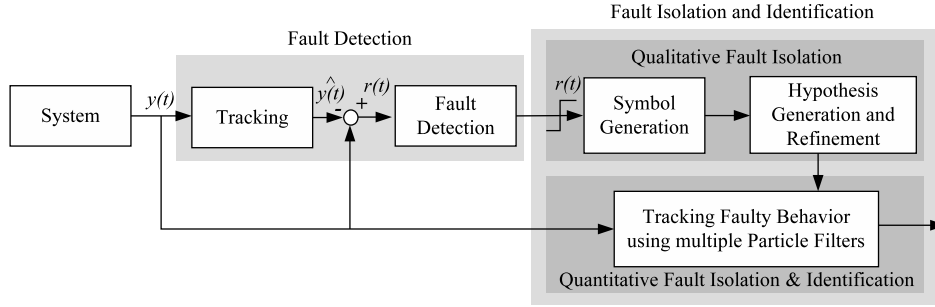


Figure 2: The diagnosis architecture.

to the observed measurements with minimum error. If, on the other hand, the measurements estimated by a PF using a particular fault deviates from the observed measurements, or the Qual-FI scheme finds this fault candidate inconsistent, we terminate the tracking of the observations using this fault model and drop this fault candidate from our list of consistent hypotheses. We discuss this approach in greater detail in the subsequent sections.

Fault Identification: Fault identification involves the determination of the magnitude or extent of a fault. Our DBN-based Quant-FII scheme combines fault isolation and identification into the same PF-based tracking process.

4 Modeling

The *bond graph* (BG) [Karnopp *et al.*, 2000] model of the system forms the core of our modeling framework. From BGs, we can systematically derive efficient models for diagnosis, the *temporal causal graphs* (TCGs) [Mosterman and Biswas, 1999] for qualitative fault isolation, and the DBNs, for detection and quantitative fault isolation and identification.

4.1 Bond Graphs

The bond graph modeling paradigm allows domain-independent, energy-based, topological modeling of physical processes. The nodes of a bond graph are energy storage (capacitors, C , and inertias, I); energy dissipation (resistors, R); energy transformation (gyrators, GY , and transformers, TF); and, energy source (sources of effort, Se , and sources of flow, Sf) elements. Nonlinear systems are modeled by parameter values that are functions of other system variables. Bonds, drawn as half arrows, represent the energy exchange pathways between the bond graph elements. Two variables, effort, e , and flow, f , are associated with each bond, and their product, $e \cdot f$ defines the rate of energy transfer through the bond. Connections in the system are modeled by two idealized elements: 0- and 1-junctions. The junctions couple two or more elements based on the principles of conservation of energy and continuity of power. Therefore, at a 0-junction (or, a 1-junction), the efforts (or, flows) of all incident bonds are equal, and the sum of flows (or, efforts) is zero. Fig. 3(b) shows the BG of a simple two-tank system (shown in Fig. 3(a)). In the hydraulics domain, a flow represents the rate of flow of fluid, and effort represents the fluid pressure.

4.2 Temporal Causal Graph

The *temporal causal graph* (TCG) structure captures the causal and temporal relations between system variables [Mosterman and Biswas, 1999]. In our work, we derive the TCG systematically from the BG model. Fig. 3(c) shows the TCG for the two-tank system. The TCG represents a signal flow graph where the effort and the flow variables in the system model are nodes, and the direction and type of interaction between variables are captured as edges. Edges are derived from component constituent relations or junction constraints. For example, an edge describing the effort-to-flow relation of a resistance, R , is labeled $1/R$, since $f = e/R$. For a capacitor in integral causality, the flow-to-effort relation is labeled dt/C , where the dt specifier implies a temporal edge, i.e., a change in the flow, f , affects the derivative of the effort, e . Junctions also impose direct (+), inverse (−), and equality (=) relations between variables.

4.3 Dynamic Bayesian Networks

A DBN is a two-slice Bayes net that not only captures the relations between system variables in any time slice t , but also captures the across-time relations between variables in time slice $t + 1$ and the previous time slice t [Murphy, 2002]. The system variables, $(\mathbf{X}, \mathbf{Z}, \mathbf{U}, \mathbf{Y})$, which represent the state variables, other hidden variables, input variables, and measured variables for the dynamic system, respectively, are all considered to be sampled from stochastic distributions. The dynamic state-space model is a discrete-time stochastic process that satisfies the first order Markov assumption. For the two-slice Bayes net, if an observed node is a function of a state variable, or an input variable, an intra-slice link is drawn from the state or input variable to that observed node. An inter-slice link, $x_t \rightarrow x'_{t+1}$, is drawn between two state variables, x_t and x'_{t+1} , if the value of x' at time $t + 1$ depends on the value of x at time t . Nodes x and x' of the DBN may represent the same state variable, but at different time points. Similarly, an inter-slice link may also be drawn between an input variable at time t , and a state variable at time $t + 1$.

The DBN model for a system can be constructed from its TCG, as outlined in [Lerner *et al.*, 2000]. First we identify the nodes, N , in the TCG that represent the state variables, system measurements, and inputs. Then for each of these nodes, $n \in N$, we create nodes n_t and n_{t+1} to denote the state of that variable at consecutive time points in the DBN. If the

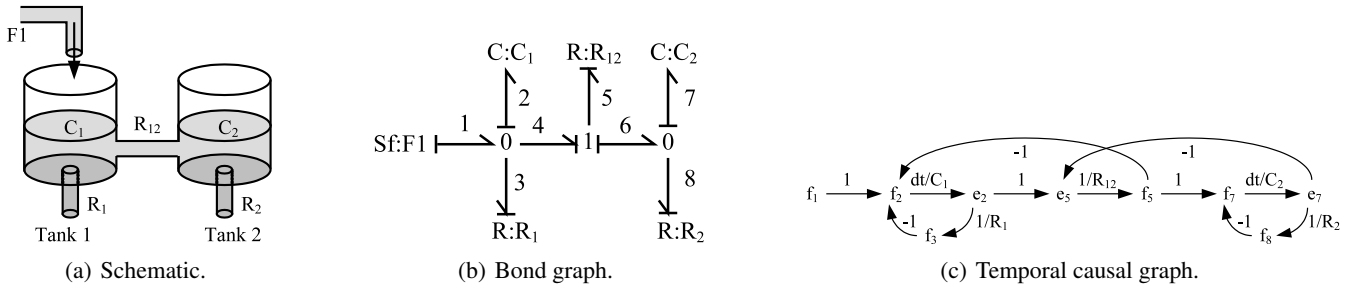


Figure 3: Two-tank system models.

relation between any two TCG nodes, $n, n' \in N$, is algebraic, links are constructed in the DBN from n_t to n'_t , and n_{t+1} to n'_{t+1} . On the other hand, if the relations between the two nodes has a delay, then a link is added in the DBN from n_t to n'_{t+1} . The DBN models for the nominal and faulty systems are explained in detail below.

Nominal System Model

The DBN for the nominal system includes nodes corresponding to state variables, observed variables, and inputs. The system component parameters are assumed to be constant, or deterministic functions defined by the nominal system model. For example, as shown in Fig. 4(a), the DBN derived from the TCG of the two-tank system has the following stochastic variables at time t : $\mathbf{X}_t = \{e_{2_t}, e_{7_t}\}$, the pressures at the bottom of tanks 1 and 2, respectively, and $\mathbf{Y}_t = \{f_{3_t}, f_{8_t}, f_{5_t}\}$, the outflows from tanks 1 and 2, and the flow between tanks 1 and 2, respectively. The input flow into tank 1, $\mathbf{U}_t = \{f_{1_t}\}$, and the component parameter values, C_1 , C_2 , R_1 , R_{12} , and R_2 are deterministic variables. $\mathbf{Z}_t = \emptyset$, i.e., the two tank dynamic model requires no additional variables.

Fault Models of the system

For each fault candidate, a separate model is derived for tracking system behavior after fault occurrence. The fault model is generated by augmenting the nominal system model with an extra state variable, representing the fault parameter. For an abrupt fault hypothesis, the fault parameter corresponds to the bias term. For an incipient fault hypothesis, the fault parameter is the drift term as well as the faulty parameter itself. This additional parameter accumulates the current value of the parameter corresponding to the incipient fault hypothesis.

The model of a two-tank system with an incipient R_1^{+i} fault includes the extra stochastic variable $\sigma_{R_1}^i$. We assume that the slope is constant, i.e., slope $\sigma_{R_1}^i(t+1) = \sigma_{R_1}^i(t)$. The fault parameter $R_1(t)$ is included as an additional stochastic variable that evolves according to the equations $R_1(t+1) = R_1(t) + \sigma_{R_1}^i(t)$, and replaces all occurrences of R_1 in the nominal model. The DBN model for this fault is shown in Fig. 4(b).

The model of a two-tank system with an abrupt R_1^{+a} fault includes the extra state variable $\sigma_{R_1}^a$. We assume that the magnitude of this bias is constant, i.e., $\sigma_{R_1}^a(t+1) = \sigma_{R_1}^a(t)$, where $t \geq t_f$. We generate the faulty system model by replacing all occurrences of R_1 in the nominal model with $(R_1 + \sigma_{R_1}^a(t))$.

Fig. 4(c) shows the DBN model for this fault.

5 Tracking and Fault Detection

The basic idea of fault detection is to track nominal system behavior using PFs, and use a statistical hypothesis testing scheme to detect statistically significant non-zero residuals.

5.1 Tracking

Particle filtering is a popular scheme for estimating the true state of a system using DBNs [Koller and Lerner, 2001]. A PF is a sequential Monte Carlo sampling method for Bayesian filtering that approximates the belief state of a system using a weighted set of samples, or particles [Arulampalam *et al.*, 2002; Koutsoukos *et al.*, 2003]. Each sample, or particle, consists of a value for each state variable, and describes a possible state the system might be in. As more observations are obtained, each particle is moved stochastically to a new state, and the weight of each particle is readjusted to reflect the likelihood of that observation given the particle's new state. The PF algorithm for DBNs is shown in Algorithm 1 [Koller and Lerner, 2001].

Algorithm 1 Particle Filtering on DBNs

Input: Number of particles, N ; a DBN $D = (\mathbf{X}, \mathbf{Z}, \mathbf{U}, \mathbf{Y})$
for each particle i , $i = 1$ to N **do**
 sample $\mathbf{X}_0^{(i)}$ from the prior $Prob(\mathbf{X}_0)$
 set $\mathbf{Y}_0^{(i)}$ to observed values at time step $t = 0$
for for each time step t **do**
 for each particle i , $i = 1$ to N **do**
 Get $[\mathbf{X}_t^{(i)}, w^{(i)}] = \text{SampleParticle}(\mathbf{Y}_t, \mathbf{X}_{t-1}^{(i)}, D)$
 Resample N new particles based on the weights $w^{(i)}$
 Generate estimated \mathbf{X}_t and \mathbf{Y}_t values from the N particles

function $\text{SampleParticle}(\mathbf{Y}_{t+1}, \mathbf{X}_t, D)$
 set $w = 1$
 for each node X in DBN D **do**
 let u be the current assignment to $\text{Parents}(X)$
 if X is not measured **then**
 sample X from $Prob(X|\text{Parents}(X) == u)$
 else if X is a measured node **then**
 set X to its observed value from \mathbf{Y}_{t+1}
 set $w = w \cdot Prob(X|\text{Parents}(X) == u)$
 return $[\mathbf{X}, w]$

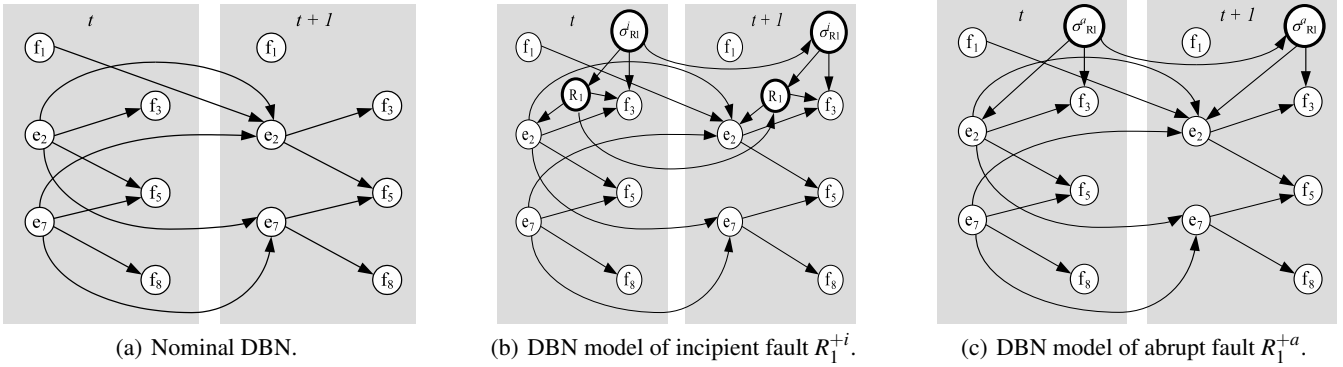


Figure 4: DBN models of a two-tank system.

We choose to run the PF scheme on the DBN models for tracking both nominal and faulty system behavior for several important reasons. Particle filtering applied to DBNs exploits the sparseness and compactness of DBNs (based on conditional independence of the variables) to provide computationally efficient solutions, especially because each observed random variable in a DBN typically depends on some, and not all state variables. The compactness of DBNs is especially noticeable in our DBN fault models, where each fault parameter typically affects the relation between a small number of state variables and measurements. Moreover, PFs are a good approximation of DBN propagation methods, when exact distributions cannot be computed analytically, especially for complex, nonlinear systems. Also, PFs can be implemented as anytime algorithms, and a trade-off between accuracy and time efficiency can be achieved by varying the number of particles [Dearden and Clancy, 2001]. Finally, the single-fault assumption allows for the decomposition of a complex multi-hypothesis isolation and identification problem into a set of simpler, single hypothesis PF-based tracking problems.

5.2 Fault Detection

The fault detector monitors each measurement residual and indicates the presence of a fault when a statistically significant non-zero fault is detected. In our work, the observer-estimated measurements are compared against the actual system measurements using a Z-test for difference in means for robust fault detection [Biswas *et al.*, 2003]. The Z-test uses a sliding window scheme to compute the residual mean and signal variance. The choice of parameters for this scheme and the confidence level chosen for the Z-test determines the properties of the fault detection filter. These parameters also determine the tradeoff between false alarms and fast detection of faults.

6 Fault Isolation and Identification

Once a fault is detected, the Qual-FI scheme is triggered to generate the initial fault hypotheses and refine these hypotheses as additional measurement deviations are observed. The Qual-FI is run till either the fault hypotheses set is refined to a pre-defined size, k , a design parameter, which is typically set

to 10% of the total number of fault hypotheses generated after a fault is detected, or a pre-specified s simulation timesteps have elapsed, after which the Quant-FII scheme is invoked to isolate and identify the true fault. We need to choose k and s carefully because if k is too large and s is too small, the large number of remaining fault candidates would make the Quant-FII inefficient. On the other hand, if k is very small, and s is large, the isolation and identification task will be delayed. In the following, we describe the two isolation schemes in more detail.

6.1 Qualitative Fault Isolation

Our qualitative fault isolation scheme is based on deriving *fault signatures* from the TCG. A fault signature is a qualitative representation of the magnitude and higher order changes in a measurement caused by a fault [Mosterman and Biswas, 1999]. The qualitative deviations of the measurements are expressed using ‘+’, ‘-’, or ‘0’ symbols which denote that the observed measurement has increased from nominal, decreased from nominal, or is nominal, respectively. Since we are dealing with noisy measurement environments, we assume that only the magnitude and slope of a signal can be reliably measured at any point in time [Manders *et al.*, 2000], and use these information to discriminate between faults. A typical fault signature of a fault f for a measurement m_1 can be $(+ -)$, which denotes a discontinuous increase followed by a gradual decrease in m_1 if fault f occurs. A $(0 -)$ signature of the same fault for another measurement, m_2 , on the other hand, implies that the occurrence of f will not generate any discontinuity in m_2 , but cause m_2 to decrease gradually.

The detection of a fault triggers the *symbol generation* module for every measurement. A sliding window scheme, similar to the one used for fault detection, is applied to the measurement residuals, and the symbols for the magnitude and slope for each measurement is determined when they deviate [Manders *et al.*, 2000].

Once a fault is detected, the hypothesis generation module of Qual-FI is invoked. This scheme propagates the changes in the parameters that are consistent with the observed deviation backwards along the TCG to generate the fault hypotheses. We combine the work reported in [Roychoudhury *et al.*, 2006] with our previous work [Mosterman and Biswas,

Table 1: Selected fault signatures for the two-tank system.

Fault	f_3	f_5	f_8
R_1^{+a}	-+	0+	0+
R_1^{+i}	0-	0+	0+
R_{12}^{+a}	0+	-+	0-
R_{12}^{+i}	0+	0-	0-
R_2^{+a}	0+	0-	-+
R_2^{+i}	0+	0-	0-

1999] to generate both abrupt and incipient fault candidates for every implicated parameter. For each abrupt fault candidate in the hypothesis set, a *forward* pass on the TCG yields the fault signatures, i.e., the effect of this fault on all remaining measurements [Mosterman and Biswas, 1999]. Since incipient faults cannot produce discontinuities in measurements, the fault signatures for incipient faults are of the form (0τ) , where τ is the first non-zero symbol in the fault signature of an abrupt fault in the same system parameter and for the same direction of change. For example, as shown in Table 1, the signature of fault R_1^{+i} for flow f_3 is $(0-)$, since that of fault R_1^{+a} for flow f_3 is $(-+)$.

As additional measurements deviate from nominal, the generated symbol deviations for these measurements are compared to the generated fault signatures, and if any fault signature is inconsistent with the observed symbol for that measurement, the fault candidate is dropped.

6.2 Quantitative Fault Isolation and Identification

Once the Qual-FI scheme discussed above refines the number of fault hypotheses to a pre-defined number, or s timesteps have elapsed, the Quant-FII scheme is started. The Quant-FII performs both fault isolation and identification. For each fault candidate that remains at the time Quant-FII is initiated, we develop a faulty DBN system model, as explained in Section 4.3. We then run a particle filter for each of these DBN fault models, taking as input the measurements from the time of fault detection, t_d , as described in Algorithm 1. As more observations are obtained, only the PF using the correct fault model, ideally, should be converging to the observed measurements, while the observations estimated by the PFs using the incorrect fault models should gradually deviate from the observed faulty measurements. A fault candidate is dropped if: (i) the Qual-FI drops that fault candidate, or (ii) the measurements estimated by that fault model significantly deviates from the observed faulty measurements.

A Z-test is used to determine if the deviation of a measurements estimated by the PF from the corresponding actual observation is statistically significant. Since even the correct fault model will need some time before the particles start converging to the observed faulty values, we need to delay the invocation of the Z-tests for s_d time steps, as otherwise, the Z-tests will indicate a deviation from observed measurements at the very onset for all fault models. We typically assume that the particles for the true fault model will converge to the observed measurements within s_d time steps of its invocation.

Since the fault magnitude is included as a stochastic variable in every fault model, the magnitude of the true fault (i.e.,

the bias, σ_p^a , or, the slope, σ_p^i) is considered to be that estimated by the PF for the true fault model.

7 Experimental Results

In this section, we present some experimental results obtained by applying the proposed diagnosis approach to the two tank system shown in Fig. 3(a). In such hydraulic systems, the accumulation of sediment in the pipes are common examples of incipient faults. In addition, sudden blockages of pipes due to the entry of foreign objects in the pipes through the tanks can be examples of abrupt faults. These incipient faults are modeled as gradual increases in pipe resistances and represented as R_1^{+i} , R_{12}^{+i} , and R_2^{+i} . Abrupt faults are modeled as step increases in the pipe resistances, and represented as R_1^{+a} , R_{12}^{+a} , and R_2^{+a} . The flows, f_3 , f_5 and f_8 , through pipes R_1 , R_{12} , and R_2 , respectively, are the measured variables for our experiments. In our experiments, we assume all random variables, and the prior and conditional probabilities are Gaussian Normal. The mean and variance of each hidden variable is set based on empirical knowledge of the model. The means and variances of the observed variables, as well as the conditional probabilities, are functions of the estimated system parameters, and the parameters of distributions of the hidden variables. For the experiments below, we set $k = 5$ and $s = 300$ s.

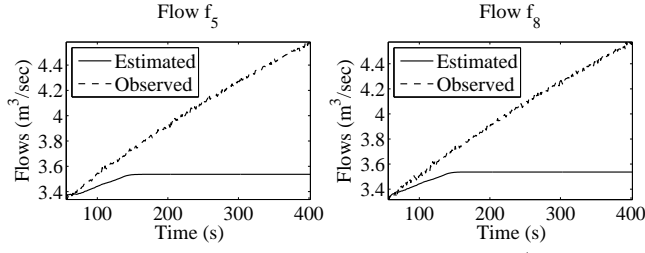
System behavior is generated for a total of 400 time steps using a Matlab Simulink simulation model. According to standard practice, white Gaussian noise with zero mean and power -40 dBW is added to the measurements. The measurements are saved to a file, and then run through our fault diagnosis scheme (implemented in Matlab) to generate our experimental results.

7.1 Experiment 1

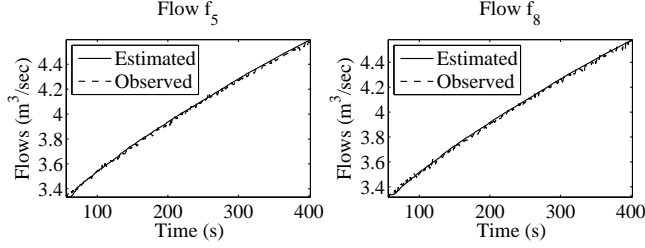
We present a run of our diagnosis scheme for a specific fault scenario. An incipient fault in pipe R_1 , R_1^{+i} , with $\sigma_{R_1}^i = 10\%$, is introduced at time step, $t = 50$ s.

For this experiment, we consider measurements f_5 and f_8 only. The R_1^{+i} causes both measurements to increase gradually from nominal. The fault detector signals an increase in f_5 at time step $t = 52$ s, followed by an increase in f_8 at time step $t = 55$ s. The symbol generator indicates that these changes are gradual, and not discontinuous. According to the fault signatures shown in Table 1, only R_1^{+a} and R_1^{+i} are consistent with the observed deviations.

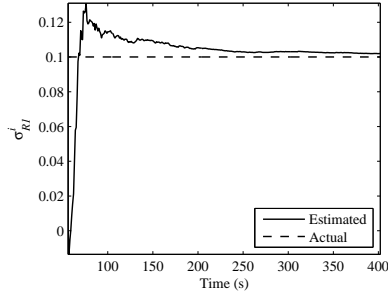
Therefore, two separate PFs, one each for R_1^{+a} and R_1^{+i} are initiated. The DBNs for the abrupt and incipient fault models are shown in Fig. 4(c) and Fig. 4(b), respectively. As more observations are obtained, the Z-tests indicate that the measurement estimates of the R_1^{+a} PF significantly deviates from the observed faulty measurements. As soon as a Z-test indicates a deviation, the only remaining fault model consistent with the observed measurements, i.e., R_1^{+i} is isolated as the true fault. While the true injected fault slope is 10%, the slope of the incipient fault, $\sigma_{R_1}^i$, is estimated to be 12% by the PF (see Fig. 5(c)). The measurements estimated by the PFs applied to the two fault models are shown in Fig. 5(a) and Fig. 5(b).



(a) Estimated observations for fault model R_1^{+a} .



(b) Estimated observations for fault model R_1^{+i} .



(c) Estimated slope of the true fault R_1^{+i} .

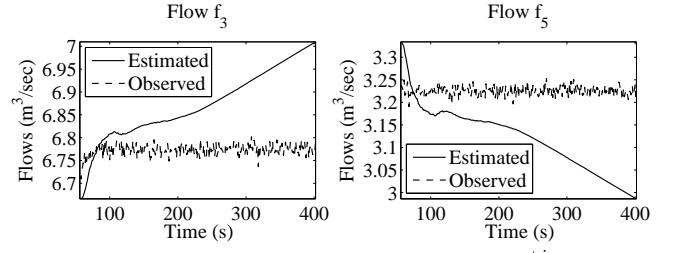
Figure 5: Experiment 1 results.

7.2 Experiment 2

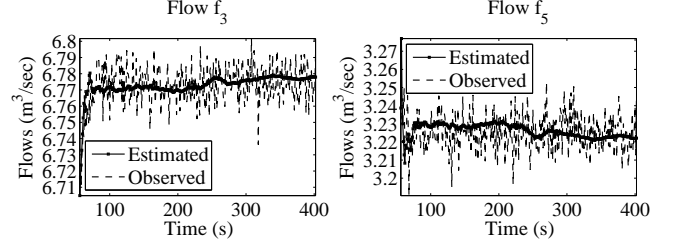
We now present another experiment for a specific fault scenario. An abrupt fault, R_2^{+a} , with $\sigma_{R_2}^a = 10\%$ is introduced in pipe R_2 at time step, $t = 50$ s. The measurements used in this experiment are flows f_3 and f_5 .

The R_2^{+a} fault causes a gradual decrease in its flow f_5 from nominal. The fault detector signals this deviation at time step $t = 52$ s. The symbol generators output the symbols ‘-’ for the slope of residual of f_5 , and ‘0’ for the residual magnitude. This is followed by a 0+ deviation in flow f_3 . As shown in Table 1, all faults but R_{12}^{+i} , R_2^{+a} , and R_2^{+i} are inconsistent with this observed decrease in magnitude of f_3 and f_8 , and hence, the fault hypothesis set is $\{R_{12}^{+i}, R_2^{+a}, R_2^{+i}\}$.

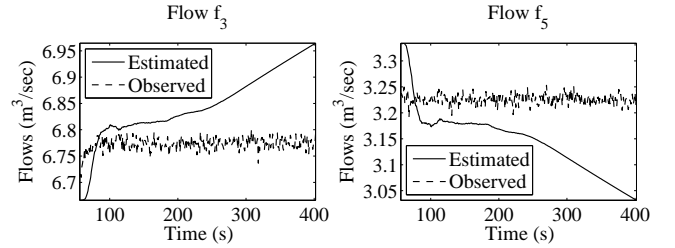
The Quant-FII scheme’s task is to both isolate and identify the magnitude of the true fault. We generate the DBN model for each of the three faults using the method described in Section 4.3, and run three PFs on these models, taking as inputs, only measurements at time points $t > 52$ s, the time of detection of the fault. Eventually, the Z-tests indicate that the observations estimated by the PFs applied to R_{12}^{+i} and R_2^{+i} have significantly deviated from the observed faulty measurements, correctly isolating fault R_2^{+a} as the true fault. The es-



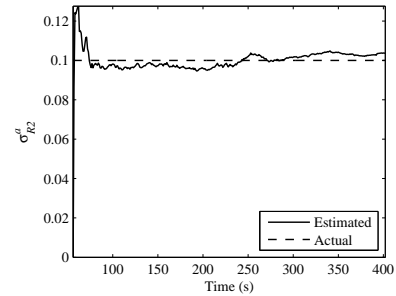
(a) Estimated observations for fault model R_{12}^{+i} .



(b) Estimated observations for fault model R_2^{+a} .



(c) Estimated observations for fault model R_2^{+i} .



(d) Estimated bias of the true fault R_2^{+a} .

Figure 6: Experiment 2 results.

timated measurements from fault models R_{12}^{+i} , R_2^{+a} , and R_2^{+i} are shown in Fig. 6(a), Fig. 6(b), and Fig. 6(c), respectively

As we can see from Fig. 6(d), the PF identifies the fault magnitude to be about a 11% step increase in R_2 , while the true fault magnitude is 10%.

8 Discussion and Conclusions

PFs have been used extensively for system health monitoring and diagnosis of hybrid systems [Dearden and Clancy, 2001; Lerner *et al.*, 2000]. The general approach involves the system to include discrete nominal and fault modes, with the evolution of the system in each discrete mode being defined using

differential equations. The process of diagnosis then involves tracking the observed measurements using a PF that runs on the comprehensive system model till the particles eventually converge to a discrete fault mode. PFs have also been used to diagnose parametric incipient and abrupt faults [Koller and Lerner, 2001]. The usual approach for using PFs for diagnosis, however, cannot alleviate the problem of *sample impoverishment*, wherein particles in faulty state (with typically very low probability, and hence low weights) are dropped during the re-sampling process. Even though several solutions to this problem have been proposed [Verma *et al.*, 2004], the diagnosis scheme still has to rank the different fault hypothesis based on their likelihoods, and report the most likely fault mode that justifies the observations the best. Our single fault assumption allows us to avoid the sample impoverishment problem by having a separate fault model for each fault hypothesis. Also, we do not rank the different fault hypotheses, and drop candidates based on their inability to track the observed faulty measurements.

In [Narasimhan *et al.*, 2004], the authors propose an approach for combining look-ahead Rao-Blackwellised PFs (RBPFs) with the consistency-based Livingstone 3 (L3) approach for diagnosing faults in hybrid systems. In this approach, the nominal RBPF-based observer tracks the system evolution till a fault is detected, after which L3 generates a set of fault candidates that are then tracked by the fault observer (another RBPF). All the fault hypotheses are included in the same model, and tracked by the fault observer. In contrast, our approach executes the qualitative and quantitative fault isolation schemes in parallel, and uses separate fault models for each fault candidate.

In the future, we seek to investigate and solve a number of open issues and problems. First, we need to study the observability of the faulty models and their impact on diagnosis. For example, in the two tank system shown in Fig. 3(a), it is not possible to uniquely discriminate between R_{12}^{+i} and R_2^{+i} faults using measurements f_3 , f_5 , and f_8 . The problem of identifying the correct set of measurements such that the system is diagnosable as well observable, therefore, is an interesting research issue. Next, we wish to apply our diagnosis approach to a large real-world system, to analyze the scalability and efficiency of our methodology. Third, we need to develop systematic procedures for obtaining the values of design parameters k , s , and s_d . Finally, we would like to improve the efficiency of our diagnosis approach by deriving reduced DBN models and running the PFs on these reduced-order models instead of on the entire system DBN model.

Acknowledgements

This work was supported in part by the National Science Foundation under Grant CNS-0615214 and NASA NRA NNX07AD12A.

References

- [Arulampalam *et al.*, 2002] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *IEEE Trans. on Signal Processing*, 50(2):174–188, 2002.
- [Biswas *et al.*, 2003] G. Biswas, G. Simon, N. Mahadevan, S. Narasimhan, J. Ramirez, and G. Karsai. A robust method for hybrid diagnosis of complex systems. In *Proc. 5th IFAC Symp on Fault Detection Supervision Safety Technical Processes*, pages 1125–1131, 2003.
- [Dearden and Clancy, 2001] R. Dearden and D. Clancy. Particle filters for real-time fault detection in planetary rovers. In *Proc. of the 12th International Workshop on Principles of Diagnosis*, pages 1–6, 2001.
- [Gertler, 1998] J. J. Gertler. *Fault Detection and Diagnosis in Engineering Systems*. Marcel Dekker, Inc., New York, NY, 1998.
- [Karnopp *et al.*, 2000] D. C. Karnopp, D. L. Margolis, and R. C. Rosenberg. *Systems Dynamics: Modeling and Simulation of Mechatronic Systems*. John Wiley & Sons, Inc., New York, NY, USA, 3rd edition, 2000.
- [Koller and Lerner, 2001] D. Koller and U. Lerner. Sampling in factored dynamic systems. In A. Doucet, N. de Freitas, and N. Gordon, editors, *Sequential Monte Carlo Methods in Practice*. Springer, 2001.
- [Koutsoukos *et al.*, 2003] X. Koutsoukos, J. Kurien, and F. Zhao. Estimation of distributed hybrid systems using particle filtering methods. In *Hybrid Systems: Computation and Control*, Lecture Notes in Computer Science, pages 298 – 313. Springer, 2003.
- [Lerner *et al.*, 2000] U. Lerner, R. Parr, D. Koller, and G. Biswas. Bayesian fault detection and diagnosis in dynamic systems. In *Proc. of Seventeenth National Conference on Artificial Intelligence*, pages 531–537, 2000.
- [Manders *et al.*, 2000] 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 on Fault Detection Supervision Safety Technical Processes*, pages 1074–1079, Budapest, Hungary, June 2000.
- [Mosterman and Biswas, 1999] P. J. Mosterman and G. Biswas. Diagnosis of continuous valued systems in transient operating regions. *IEEE-SMCA*, 29(6):554–565, 1999.
- [Murphy, 2002] K. P. Murphy. *Dynamic Bayesian Networks: Representation, Inference, and Learning*. PhD thesis, University of California, Berkeley, 2002.
- [Narasimhan *et al.*, 2004] S. Narasimhan, R. Dearden, and E. Benazera. Combining particle filters and consistency-based approaches for monitoring and diagnosis of stochastic hybrid systems. In *Proc. of the 15th International Workshop on Principles of Diagnosis*, 2004.
- [Roychoudhury *et al.*, 2006] I. Roychoudhury, G. Biswas, and X. Koutsoukos. A bayesian approach to efficient diagnosis of incipient faults. In *Proc. of the 17th International Workshop on Principles of Diagnosis*, pages 243–250, Spain, 2006.
- [Verma *et al.*, 2004] V. Verma, G. Gordon, R. Simmons, and S. Thrun. Real-time fault diagnosis. *Robotics & Automation Magazine, IEEE*, 11(2):56–66, 2004.