

Distributed Diagnosis of Dynamic Systems Using Dynamic Bayesian Networks^{*}

Indranil Roychoudhury Gautam Biswas Xenofon Koutsoukos

*Institute for Software Integrated Systems, Department of EECS,
Vanderbilt University, Nashville, TN, USA 37235,
{indranil.roychoudhury, gautam.biswas, xenofon.koutsoukos}@vanderbilt.edu*

Abstract: This paper presents a Dynamic Bayesian Network (DBN)-based distributed diagnosis scheme, where each distributed diagnoser generates globally correct diagnosis results without a centralized coordinator by communicating a minimal number of measurements so that each diagnoser satisfies local observability properties, and the overall diagnoser is globally observable. We present a procedure for designing the distributed diagnosers by factoring a system DBN into the maximal number of smaller DBN Factors (DBN-Fs) that are conditionally independent of other DBN-Fs, given the communicated measurements. Since each conditionally independent DBN-F is observable, Bayesian inference schemes can be applied to each factor *independently* for distributed tracking of system behavior for isolation and identification of faults without loss of accuracy. We prove that each local diagnoser guarantees globally correct diagnosis results, and present some experimental results for an electrical circuit to demonstrate the efficacy of our diagnosis scheme.

1. INTRODUCTION

To ensure safe and efficient operation of real-world engineering systems, online model-based diagnosis schemes must be robust to uncertainties, such as sensor noise and modeling inaccuracies. Dynamic Bayesian Networks (DBNs) provide a systematic method for modeling the behavior of dynamic systems in uncertain environments (Murphy [2002]). A DBN is a directed acyclic graph structure that represents a probabilistic discrete-time model of a dynamic system. Nodes in the graph represent random variables, and links denote causal dependencies between nodes within a time step, and across time steps. DBNs exploit the conditional independence between system variables to provide a compact representation for reasoning about dynamic systems behavior. Bayesian inference algorithms have been widely used for diagnosis of dynamic systems represented as DBNs (e.g., see Lerner et al. [2000] and Roychoudhury et al. [2008]). Unfortunately these centralized schemes are expensive in memory and computational requirements, scale poorly to changes in system configuration, and have single points of failure. Distributed diagnosis schemes can address these drawbacks of centralized schemes, as shown in Pencole and Cordier [2005], Debouk et al. [2000] and Roychoudhury et al. [2009a].

This paper presents a DBN-based distributed diagnosis approach, where each distributed diagnoser generates globally correct diagnosis results by communicating a minimal number of measurements with each other, and not requiring a centralized coordinator. The notion of structural observability applied to bond graph (BG) models (Seur and Dauphin-Tanguy [1991]) is exploited to derive DBN factors (DBN-Fs) that are independently observable, and

together retain observability for the entire system. We have developed systematic methods for deriving the DBN-Fs from a BG model of the system (Roychoudhury et al. [2009b]), and in this paper, we prove that these factors can be used as local diagnosers that generate globally correct results without loss of accuracy.

We implement a particle filter (PF)-based (see Koller and Lerner [2001]) inference approach on each DBN-F for fault detection, isolation and identification, and employ a qualitative fault isolation scheme to improve diagnosis efficiency. We prove that distributed diagnosers do not need a coordinator for generating globally correct diagnosis results, since the effects of a fault in one DBN-F propagates to other factors only through the communicated measurements, which are now considered inputs to the different diagnosers. Therefore, the PFs that are implemented for the other remote DBN-Fs track the faulty data without detecting a fault, and the isolation schemes in the remote diagnosers do not get activated.

This paper is organized as follows. Section 2 presents background on modeling for diagnosis, and our diagnosis approach. Section 3 presents our diagnoser design approach based on factoring the DBNs into conditionally independent and observable DBN-Fs, as well as, our distributed diagnosis architecture. Section 3 also discusses the important properties of our distributed diagnosis scheme. Section 4 presents the results of diagnosis experiments on an electrical circuit. Section 5 presents related work, and Section 6 concludes the paper.

2. BACKGROUND

2.1 Modeling for Diagnosis

In our work, we systematically derive the diagnosis models for fault isolation in the form of *temporal causal graphs*

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

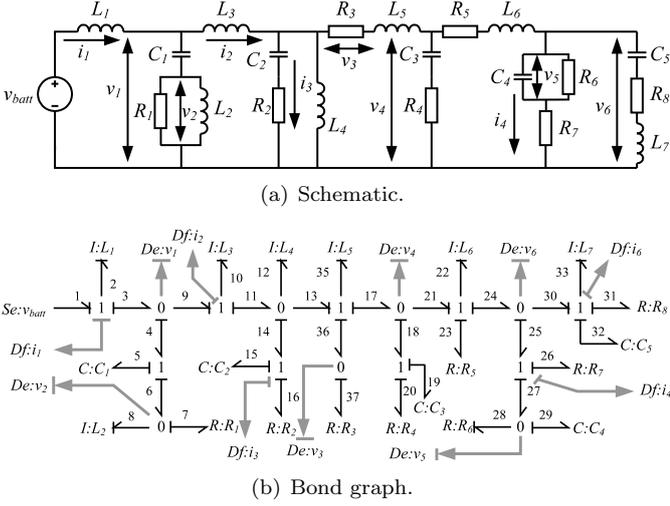


Fig. 1. Electrical circuit models.

(TCGs) (see Mosterman and Biswas [1999]), and DBNs for fault detection and identification (see Roychoudhury et al. [2008]). All of these models are derived from the system's *bond graph* (BG) model (see Karnopp et al. [2000]).

BGs are topological models that capture energy exchange pathways in physical processes. The generic elements in BGs are energy storage (C and I), dissipation (R), transformation (GY and TF), source (Se and Sf), and detection (De and Df) elements. The connecting edges, called *bonds*, represent energy pathways between the elements. Each bond, numbered i , has an associated effort, e_i , and flow, f_i , variable, such that $e_i \times f_i$ defines the power transferred through the bond. 0- and 1-junctions represent parallel and series connections, respectively. Fig. 1(b) shows the BG of a twelfth-order electrical circuit shown in Fig. 1(a). In the electrical domain, the effort variables denote voltage difference across, and flow variables denote current through, BG elements. For example, $f_2 = i_1$ denotes the current through the inductor L_1 , and $e_7 = v_2$ denotes the voltage difference across resistor R_1 . $e_1 = v_{batt}$ denotes the voltage imposed by the voltage supply. $De : v_2$ is a voltage sensor.

A TCG is essentially a signal flow graph that captures the causal and temporal relations between its nodes, which represent system variables, through directed *edges* and their *labels*. The direction of a TCG edge and its label are based on *causality*, which establishes the cause and effect relationships between the e_i and f_i variables of a bond i based on constraints imposed by the incident BG elements. The sequential causal assignment procedure (SCAP) systematically assigns the causality in a BG (see Karnopp et al. [2000]). Energy storage elements can either impose *integral* (preferred) or *derivative* causality. For example, for a C element in integral causality, $e_i = (1/C) \int f_i dt$, and hence the TCG shows $f_i \xrightarrow{dt/C} e_i$, with dt denoting a temporal relationship between f_i and e_i . For a C element in derivative causality, the corresponding TCG edge is $e_i \xrightarrow{C/dt} f_i$, since $f_i = C de_i/dt$.

A DBN can be defined as $\mathcal{D} = (\mathbf{X}, \mathbf{U}, \mathbf{Y})$, where \mathbf{X} , \mathbf{U} , and \mathbf{Y} are sets of stochastic random variables that

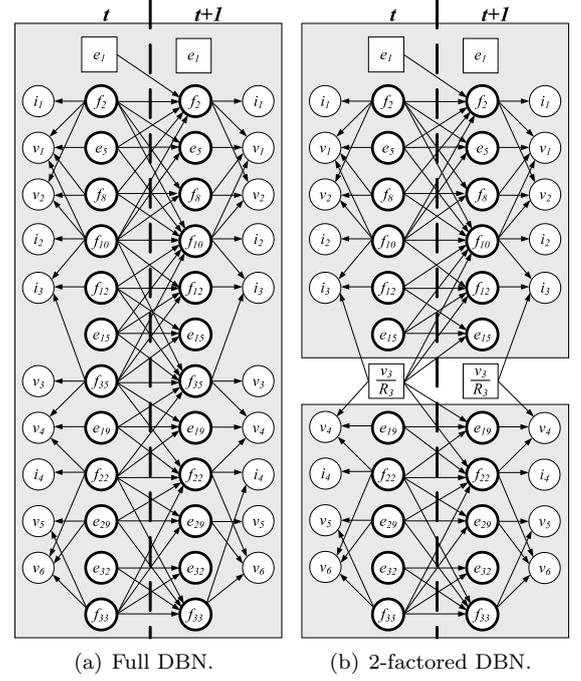


Fig. 2. DBNs for the electrical circuit.

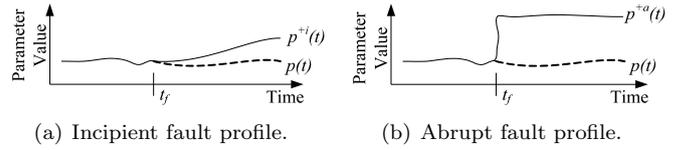


Fig. 3. Fault profiles.

denote (hidden) state variables, system input variables, and measured variables in the dynamic system, respectively (see Murphy [2002]). Graphically, a DBN is a two-slice Bayesian network, representing a snapshot of system behavior in two consecutive time slices, t and $t + 1$. Each DBN time-slice represents the Markov process observation model, $P(\mathbf{Y}_t | \mathbf{X}_t, \mathbf{U}_t)$, while the across-time links represent the Markov state-transition model, $P(\mathbf{X}_{t+1} | \mathbf{X}_t, \mathbf{U}_t)$. The system DBN is constructed from its TCG in integral causality using the method given in Lerner et al. [2000]. Fig. 2(a) shows the DBN for our example circuit, where thick-lined circles denote state variables, thin-lined circles denote observed variables, and squares denote input variables.

2.2 Modeling Faults

Our diagnosis scheme is geared toward isolating and identifying incipient and abrupt faults in discrete-time continuous dynamic systems. An *incipient fault* is a slow change in a system parameter, p (with nominal parameter value function, $p(t)$), and modeled as a linear function with a constant slope, σ_p^i , added to the nominal component parameter value function, $p(t)$, i.e., $p^{\pm i}(t) = p(t) \pm \sigma_p^i \times (t - t_f)$, $t > t_f$, where t_f is the time of fault occurrence, and $p^{\pm i}(t)$ is the temporal profile of parameter p with an incipient fault. An *abrupt fault* is modeled as an addition of a constant persistent bias term, σ_p^a , to the nominal parameter value, $p(t)$, i.e., $p^{\pm a}(t) = p(t) \pm \sigma_p^a$, $t > t_f$, where t_f is the time of fault

occurrence, and $p^{\pm a}(t)$ is the temporal profile of parameter p with an abrupt fault. Fig. 3 shows an incipient and an abrupt fault profile.

2.3 Our Diagnosis Approach

Our combined qualitative-quantitative model-based diagnosis approach was introduced in Roychoudhury et al. [2008], and has three primary components: (i) fault detection, (ii) qualitative fault isolation (Qual-FI), and (iii) fault hypothesis refinement and identification (FHRI). In the following, we present the diagnosis approach briefly, and refer the reader to Roychoudhury et al. [2008] for details. As shown in Fig. 4, each individual distributed diagnoser performs diagnosis using this approach.

Fault Detection For fault detection, a PF-based observer is implemented on the *nominal* DBN-F for each diagnoser to track nominal system behavior. In a “nominal” DBN-F, only the state and measurement variables are considered as random variables, and the system parameters are considered to be deterministic. 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 (see Koller and Lerner [2001]). Each sample, or particle, consists of a value for each state variable, and describes a possible system state. 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. A fault is *detected* when the difference between the observed (faulty) and estimated (nominal) values of any measurement is determined to be statistically significant using a statistical Z-test (see Manders et al. [2000]), having accommodated for measurement noise and modeling error.

Qualitative Fault Isolation Once a fault is detected, the *symbol generator* of Qual-FI is activated, which uses a sliding window scheme to express the magnitude and slope of *every* measurement as qualitative ‘+’, ‘−’, or ‘0’ symbols, denoting that the observed measurement has increased from nominal, decreased from nominal, or is at nominal, respectively (see Manders et al. [2000]). In the meanwhile, the *hypotheses generation* module propagates the first observed measurement-deviation backwards along the TCG, and identifies the set of all possible parameter changes that explain the observed deviation. As explained in Roychoudhury et al. [2008], we generate both abrupt and incipient fault hypotheses.

The fault hypotheses are refined by comparing the *fault signatures* of the fault hypotheses, i.e., the qualitative representation of the magnitude and higher order changes in a measurement caused by a fault and expressed as qualitative ‘+’, ‘−’, and ‘0’ symbols (Mosterman and Biswas [1999]), to the observed measurement deviations, and dropping fault hypotheses inconsistent with the observed deviations from consideration. Fault signatures are generated from the system TCG. Example fault signature of a fault, say p^{+a} for a measurement m_1 can be (+−), denoting a discontinuous increase followed by a gradual decrease in m_1 if fault p^{+a} occurs, or (0−), denoting a gradual decrease in m_2 when p^{+a} occurs.

The Qual-FI scheme is run till either the fault hypotheses set is refined to a pre-defined size, k , a design parameter, or a pre-specified s simulation timesteps have elapsed, after which the FHRI scheme is invoked to isolate and identify the true fault.

Fault Hypothesis Refinement and Identification The FHRI performs both fault hypothesis refinement and identification if multiple fault hypotheses remain when FHRI is initiated. If however, the Qual-FI has refined the set of hypotheses to a singleton, FHRI performs the task of fault identification only. For each fault hypothesis that remains at the time FHRI is initiated, a faulty system model is generated by extending the nominal DBN-F to include the fault parameter as a stochastic variable in the DBN-F, as explained in Roychoudhury et al. [2008]. A PF approach is then implemented using each DBN-F fault model, taking as input the measurements from the time of fault detection, t_d , to track the faulty behavior. As more observations are obtained, ideally the PF using the correct fault model will converge to the observed measurements, while the observations estimated using the incorrect fault models should gradually deviate from the observed measurements. A fault hypothesis is removed from consideration 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 measurement 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 measurement 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.

The specific problem we are trying to solve in FHRI is a combined parameter and state estimation problem, where we consider the otherwise “constant” fault variable as part of an extended state vector. As a result, our FHRI approach is prone to the usual “particle attrition” and “weight degeneracy” problems, as discussed in Liu and West [2000]. In this paper, we adopt the location shrinkage-based solution presented in Liu and West [2000] wherein a “shrinking” or decaying variance is added to the fault variable to ensure that enough samples of the fault variable are generated near its actual true mean, and particle attrition is avoided.

3. DISTRIBUTED DIAGNOSIS ARCHITECTURE

The basis of our distributed diagnosis approach is construction of the local diagnosers from *observable* DBN-Fs that are conditionally independent. A system is observable if the hidden states of the system can be unambiguously determined based on the observed measurements. The

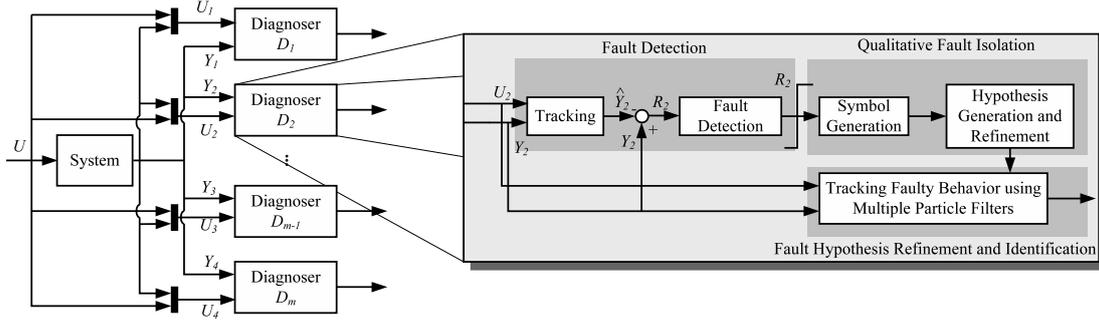


Fig. 4. The distributed diagnosis architecture.

observability of DBN-Fs permit our factored inference scheme to generate accurate inference results. The rest of this section discusses the observability property and the design of conditionally independent observable DBN-Fs, presents our distributed diagnosis approach, and provides a proof of how the design of the distributed diagnosers ensures that the local diagnosers generate globally correct diagnosis without a centralized coordinator.

3.1 Designing the Distributed Diagnosers

The objective of our distributed diagnosis scheme is to generate globally correct diagnosis results without a centralized coordinator, and by communicating a minimal number of measurements between diagnosers. We achieve this objective by factoring the system DBN, $\mathcal{D} = (\mathbf{X}, \mathbf{U}, \mathbf{Y})$, into *maximal* number of conditionally independent *DBN Factors* (DBN-Fs), $\mathcal{D}_i = (\mathbf{X}_i, \mathbf{U}_i, \mathbf{Y}_i)$, $i \in [1, m]$, such that each DBN-F is *observable*.

Definition 1. (DBN Factor). A *DBN Factor* (DBN-F), $\mathcal{D}_i = (\mathbf{X}_i, \mathbf{U}_i, \mathbf{Y}_i)$, $i \in [1, m]$, of DBN $\mathcal{D} = (\mathbf{X}, \mathbf{U}, \mathbf{Y})$ is a smaller DBN such that (i) $\bigcup \mathbf{X}_i \subset \mathbf{X}$, (ii) $\bigcup \mathbf{Y}_i \subset \mathbf{Y}$, (iii) $\bigcup \mathbf{U}_i = \mathbf{U} \cup (\mathbf{Y} - \bigcup \mathbf{Y}_i)$, and (iv) each \mathcal{D}_i is *conditionally independent* from all other DBN-Fs given the inputs, \mathbf{U}_i .

A DBN-F, \mathcal{D}_j , is termed *conditionally independent* of other DBN-Fs, \mathcal{D}_k ($k \neq j$), given its inputs, \mathbf{U}_j , if every random variable in \mathcal{D}_j is conditionally independent of all other variables in \mathcal{D}_k given \mathbf{U}_j .

Observability of each DBN-F is crucial for our monitoring and diagnosis application to ensure efficient and accurate tracking of nominal system behavior when a PF algorithm is applied to each DBN-F separately. We term a DBN-F \mathcal{D}_j to be *observable* if the underlying subsystem it represents is *structurally observable* (see Seur and Dauphin-Tanguy [1991]). Unlike previous factoring schemes, such as Boyen and Koller [1998] and Ng and Peshkin [2002], our factoring scheme preserves the system dynamics, and *does not* approximate the belief state. Hence, as shown in Roychoudhury et al. [2009b], our factored inference scheme improves efficiency of estimation without sacrificing accuracy of estimation much.

Our factoring procedure is briefly described below. Details of this factoring approach, and related formal derivations and proofs can be found in Roychoudhury et al. [2009b] and Roychoudhury et al. [2009c], respectively. Our procedure for *factoring* a DBN involves replacing one or more of its state variables by algebraic functions of at most r

measured variables, \mathbf{Y}^r , where r is a user-specified parameter. Once we express a state variable in terms of \mathbf{Y}^r , i.e., $X = g^{-1}(\mathbf{Y}^r)$, considering \mathbf{Y}^r to be inputs, we delete every $X_t \rightarrow X_{t+1}$, $U_t \rightarrow X_{t+1}$, $X_t \rightarrow Y_t$ link, and replace X with $g^{-1}(\mathbf{Y}^r)$. Then, we restore an intra-time slice link $g^{-1}(\mathbf{Y}^r) \rightarrow Y_t$ for every $X_t \rightarrow Y_t$, such that $Y_t \notin \mathbf{Y}^r$. The across-time links into X_t are not restored, since $g^{-1}(\mathbf{Y}^r)$ can be computed independently at each time step. The replacing of sufficient number of state variables in terms of measurements, and the subsequent removal of across-time links involving these state variables produces conditionally independent DBN-Fs.

Fig. 2(b) shows the DBN of the electrical circuit factored into two DBN-Fs. We assume $r = 1$ in the following. It is evident from Fig. 1(a) that the current through the inductor L_5 is equal to v_3/R_3 . Hence, we can replace f_{35} in Fig. 2(a) with v_3/R_3 , as shown in Fig. 2(b). Since, v_3/R_3 can be *measured* at every time step, all causal links into this node is removed. As a result, given v_3/R_3 , every variable in one factor is conditionally independent of the variables in the other factor. Thus, two conditionally independent factors are generated.

There are two situations in which a state variable is not removed from a global DBN: (i) if the removal of this state variable does not generate any new factors, e.g., the state variables f_2 and f_{33} are not replaced by functions of i_1 and i_6 , respectively, as that would not generate any more factors in Fig. 2(b), and (ii) if the state variable is associated with an energy storage element that is assumed to be a possible fault candidate, e.g., the state variable f_{10} is not replaced because we assume inductor L_3 can have faults, and hence need to be retained in faulty DBN models.

We generate the maximum number of *observable* DBN-Fs from a given system DBN using a two-step procedure: (i) generate maximal number of factors possible by replacing every state variable which can be determined as a algebraic function of at most r measurements, and (ii) *merge* unobservable DBN-Fs from this maximal factoring into other factors till all of the generated factors are observable. Since DBNs can systematically derived from BGs, the structural analysis of the BG fragment (BG-F) representing a DBN-F can determine if the system is *structurally observable*, as described in Seur and Dauphin-Tanguy [1991]. A system is structurally observable if in its BG, (i) there exists at least one causal path for each I and C element in the preferred integral causality to a sensor element De or Df , and (ii)

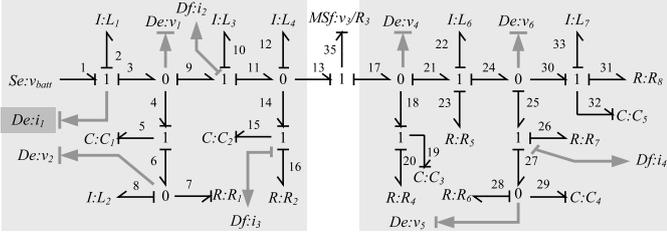


Fig. 5. Two-Factored circuit bond graph with imposed derivative causality.

inverting the causality of *every* I and C element initially in integral (preferred) causality still produces a valid causal assignment for the entire BG¹.

Given a DBN-F D_i , we can test whether or not it is observable by first mapping D_i to a BG-F, and analyzing this BG-F, B_i for structural observability. Before mapping a D_i to a B_i , we identify the state variables in the global DBN that were removed to generate D_i , and the measurement variables these state variables were replaced with. Given this information, the first step of mapping a D_i to a B_i is to replace the I or C element (in the global BG) corresponding to each state variable that was removed from the global DBN to generate D_i by a Sf or Se element, respectively, whose value is computed in terms of at most r measurements. Then, we define B_i to be that fragment of the system BG that lies between these newly introduced Sf or Se elements, as the BG is factored into independent subsystems by these source elements. We can see that the DBN-Fs shown in Fig. 2(b) map to the BG-Fs shown in Fig. 5. Both the BG-Fs are structurally observable as they fulfill both the conditions necessary for structural observability mentioned above. Note that the current sensor i_1 had to be dualized to assign derivative causality to the BG-F on the left in Fig. 5.

We propose *merging* of two or more unobservable DBN-Fs to generate an observable DBN-F. k DBN-Fs, $\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_k$, can be merged by restoring those state variables in the system DBN that were replaced to generate $\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_k$, redrawing the across-time causal links involving these state variables, and reintroducing the measurements that were used to compute these state variables. For details, please see Roychoudhury et al. [2009c]. Since the two BG-Fs shown in Fig. 5 are structurally observable, we require any further merging in our particular example.

Once a system DBN is factored into m DBN-Fs, $\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_m$, we construct a distributed diagnoser, D_i , based on each DBN-F \mathcal{D}_i . A diagnoser D_i is responsible for diagnosing faults F_i based on its observations \mathbf{U}_i .

3.2 Distributed Diagnosis Scheme

The distributed diagnosis architecture is shown in Fig. 4. Each distributed diagnoser D_i receives input signals \mathbf{U}_i , and observed measurements \mathbf{Y}_i from the system. Note that a diagnoser D_i 's inputs, \mathbf{U}_i , may include some of the inputs to the global system, i.e., $\mathbf{U}_i \cap \mathbf{U} \neq \emptyset$, as well as some measurements now considered inputs, i.e., $\mathbf{U}_i \cap \mathbf{Y} \neq \emptyset$. Two diagnosers D_j, D_k *communicate* a measurement

¹ In some situations, this may require changing a De or Df element into their dual form

$Y \in \mathbf{Y}$ if $Y \in \mathbf{U}_j \wedge Y \in \mathbf{U}_k$, i.e., measurement Y is an input to both D_j and D_k .

Each diagnoser D_i implements a distributed PF-based observer on its DBN-F \mathcal{D}_i . Because the DBN-Fs are conditionally independent, we can implement a PF on each DBN-F as an independent process. Each of these PFs takes as inputs, \mathbf{U}_i , and estimates \mathbf{X}_i based on \mathbf{Y}_i . The particle filters only communicate measurements $(\cup_i \mathbf{U}_i) - \mathbf{U}$ between themselves. The PF for the DBN-F \mathcal{D}_i uses $a \frac{|\mathbf{X}_i|}{|\mathbf{X}|}$ particles, where a is a user-specified parameter. Given m DBN-Fs, we know that $\sum_i |\mathbf{X}_i| < |\mathbf{X}|$, where \mathbf{X} is the total number of state states in the complete system. Therefore, the complexity of tracking using each DBN-F is less than that of tracking using the global DBN. Also, since the inference algorithms on the different factors are executed simultaneously, the total complexity of inference reduces to the complexity of inference of the particle filter with the maximum number of particles.

As explained in Section 2.3, each of the distributed PFs can be used on the nominal DBN-F \mathcal{D}_i for tracking nominal system behavior, and detecting faults. Qual-FI is performed using the measurements in each \mathcal{D}_i , and FHRI involves including each \mathcal{D}_i by including fault variables as extra state variables.

In our approach, we assume single, persistent, parametric faults. We start by estimating the nominal behavior of state variables in each factor by running PF-based observers in parallel. The observers can be run independently of one another due to the independence of a factor, guaranteed by construction. This independent execution of the observers in each diagnoser results in the following property.

Property 1. The failure of one of the observers will not affect the quality of state estimates at other observers.

Once a fault in F_j is detected in any one diagnoser D_j , as explained in Section 2.3, first the Qual-FI is initiated, followed by Quant-FII, till the true fault is diagnosed. Given the way the DBN-Fs are constructed, we can argue that our distributed diagnosers fulfill the following property.

Property 2. A fault $\phi \in F_j$ can be detected by diagnoser D_j only, and all other diagnosers, $D_k, k \neq j$, will not detect the fault, and hence not get activated, even though the effect of fault ϕ propagates to all other factors.

Proof: From Section 3.1, we know that every DBN-F \mathcal{D}_i has a one-to-one mapping to a BG-F B_i . As a running example, note that the two DBN-Fs (say \mathcal{D}_1 and \mathcal{D}_2) shown in Fig. 2(b) correspond to the BG-Fs, B_1 and B_2 shown in Fig. 5. A diagnoser D_i is activated only when a fault is detected by it. In general, let us assume that the observer in diagnoser D_i uses the state space equations $\hat{\mathbf{X}}_{i,t+1} = G_i(\mathbf{X}_{i,t}, \mathbf{U}_{i,t})$, and $\hat{\mathbf{Y}}_{i,t} = H_i(\mathbf{X}_{i,t}, \mathbf{U}_{i,t})$. Let us now assume that there is a fault in BG-F B_k . This means that functions G_k and H_k do not correctly represent the actual system any more. As a result, $\hat{\mathbf{Y}}_k \neq \mathbf{Y}_k$, and a fault is eventually detected by D_k . The effects of a fault in B_k can propagate to another BG-F $B_j, j \neq k$, through the shared inputs, $(\mathbf{U}_j \cap \mathbf{U}_k) - \mathbf{U}$, iff B_k and B_j communicate at least one measurement, i.e., $(\mathbf{U}_k \cap \mathbf{U}_j) - \mathbf{U} \neq \emptyset$. But,

Table 1. Fault Signatures for Diagnoser D_1

Fault	i_3	i_1	i_2	v_1	v_2
$C_2^{-a}, C_2^{-i}, R_2^{+a}, R_2^{+i}$	0-	0-	0-	0+	0+
L_2^{-a}	0-	0+	0-	-+	-+
L_2^{-i}	0-	0+	0-	0-	0-
L_3^{-a}	+ -	0+	+ -	-*	-+
L_2^{+i}	0+	0+	0+	0-	0-
L_3^{+a}	-+	0+	0+	0-	0-
L_4^{-i}	0-	0+	0+	0-	0-

Table 2. Fault Signatures for Diagnoser D_2

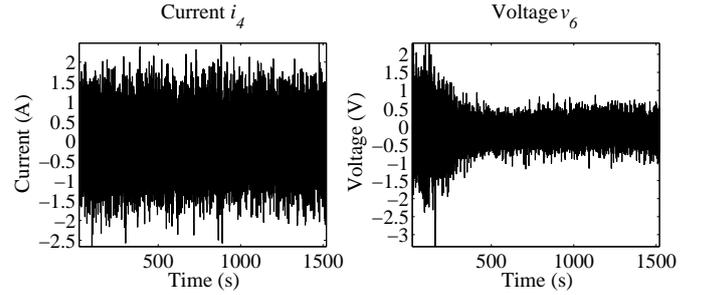
Fault	i_4	v_6	v_4	v_5
C_3^{-a}, R_4^{+a}	0+	0+	+ -	0+
C_3^{-i}, R_4^{+i}	0+	0+	0+	0+
C_4^{-a}	0-	+ -	0+	+ -
$C_4^{-i}, R_6^{+a}, R_6^{+i}$	0-	0+	0+	0+
L_7^{-a}	-+	-*	0-	0-
L_7^{-i}	0-	0-	0-	0-
R_7^{+a}, R_7^{+i}	0-	0+	0+	0-

since we adopt the single-fault assumption, and since by construction, two BG-Fs can never share any parameter, the state space representations G_j and H_j of all other BG-Fs, B_j , $j \neq k$, will correctly represent the actual system dynamics of each BG-F. Hence, $\hat{\mathbf{Y}}_j \approx \mathbf{Y}_j$, i.e., the observers in other diagnosers will correctly track the faulty measurement, and hence no fault will be detected. Consequently, if a fault is not detected, the diagnoser will not be activated.

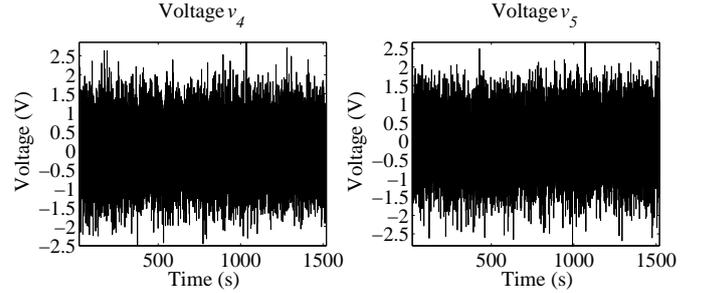
4. EXPERIMENTAL RESULTS

In this section, we present experimental results obtained by applying the proposed distributed diagnosis approach to the electrical circuit shown in Fig. 1(a). Two distributed diagnosers D_1 and D_2 are designed for this electrical circuit, for the top and bottom DBN-F shown in Fig. 2(b), respectively. Usual faults in such electrical circuits include degradation of capacitors and inductors, and increase in resistances. As explained earlier, the global DBN of the circuit can be factored into two DBN-Fs, shown in Fig. 2, and a distributed diagnoser is constructed from each DBN-F. The two diagnosers communicate measurement v_3 between each other. Tables 1 and 2 show the possible faults that must be diagnosed by each of the two diagnosers, and the qualitative fault signatures for each fault, given the measurements available to each diagnoser.

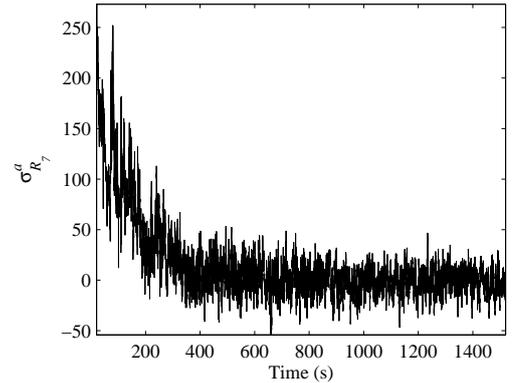
In our experiments, we assumed all random variables to be sampled from Gaussian Normal distributions. The mean and variance of each hidden variable was set based on empirical knowledge of the system and sensors. The means and variances of the observed variables, as well as the conditional probabilities, are functions of the estimated system parameters, and the parameters of distributions of the hidden variables. For the experiments below, we set $k = 2$ and $s = 300$ s. System behavior was generated for a total of 800 time steps using a Matlab Simulink simulation model. Gaussian white noise with zero mean and power -3.01 dBW was added to all measurements.



(a) Estimation errors for i_4 and v_6 for fault model R_7^{+a} .



(b) Estimation errors for v_4 and v_5 for fault model R_7^{+a} .



(c) Estimation error for $\sigma_{R_7}^a$ for fault model R_7^{+a} .

Fig. 6. Experimental 1 results.

4.1 Experiment 1

We present a run of our diagnosis scheme for a specific fault scenario. An abrupt fault in R_7 , R_7^{+a} , with $\sigma_{R_7}^a = 250$, is introduced at time step, $t = 20$ s.

From Table 2, we can see that R_7^{+a} causes a gradual decrease in i_4 and v_5 , and gradual increase in v_4 and v_6 from the point of fault occurrence. The fault detector first detects a 0- change in i_4 , and hence, the Qual-FI generates $C_4^{-i}, R_6^{+a}, R_6^{+i}, L_7^{-i}, R_7^{+a}$, and R_7^{+i} as possible fault hypotheses which could explain the observed 0- change in i_4 . Then, when a 0- change in v_5 is observed, the fault hypotheses are refined to L_7^{-i}, R_7^{+a} , and R_7^{+i} . After this, L_7^{-i} is dropped from consideration when a 0+ deviation is observed in measurement v_4 . Table 2 shows that R_7^{+a} and R_7^{+i} cannot be discriminated qualitatively, and since $k = 2$, the Quan-FII is initiated.

Two separate PFs, one each for R_7^{+a} and R_7^{+i} are initiated. As more observations are obtained, the Z-tests indicate that the measurement estimates of the R_7^{+i} PF signifi-

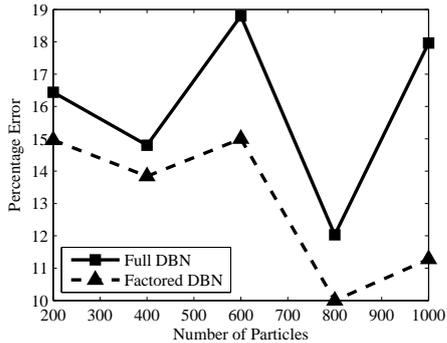


Fig. 7. Percentage error in estimation of $\sigma_{R_7}^a$.

cantly 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_7^{+a} is isolated as the true fault. It can be seen from Fig. 6(c) that the estimated fault magnitude converges to the actual magnitude of the R_7^{+a} fault that was introduced. The estimation errors of the PF applied to the abrupt fault model is shown in Figs. 6(a) and 6(b). As expected, diagnoser D_1 observer tracks the system observations without detecting any measurement deviation, and hence, activating the Qual-FI in D_1 .

4.2 Experiment 2

The following experiment demonstrates that our distributed diagnosis scheme does not sacrifice accuracy for improvement of efficiency. To demonstrate this, we generated two fault models for fault R_7^{+a} , one using the global DBN, and the other using a DBN-F, and ran a PF to identify the true fault magnitude using the two fault models, with increasing number of particles. Fig. 7 shows the percentage error in identifying the true fault magnitude for the PF using the full DBN and the factored DBN, and Fig. 8 shows the time each PF took to converge to within 20% of the true fault magnitude. The results show that for the same number of particles, our distributed FHRI scheme is more accurate, as well as, efficient than the centralized approach. Also the increase in time taken as the number of particles are increased occurs at a slower rate for the factored DBNs. This is expected because the global DBN has about double the number of state variables than the DBN-F. However, in Section 3, we described how in our distributed diagnosis scheme, the total number of particles available are proportionally distributed amongst the PFs implemented on different DBN-Fs based on the number of hidden variables in each DBN-F. Given this scheme, we can see that a PF on a DBN-F using 200 particles gives more accurate estimates than a PF on the global DBN with 400 particles, and so on, in less time. Hence, we validate that our distributed diagnosis scheme does not sacrifice accuracy for improved efficiency.

5. RELATED WORK

Decentralized diagnosis schemes can be broadly classified to conform to one of the three protocols presented in Debouk et al. [2000], where each local diagnoser is built from the global system model and uses only a subset of observable events. Coordination is necessary in the first and

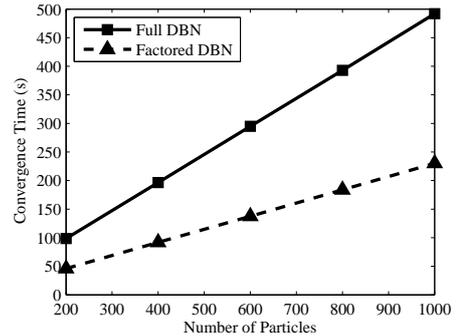


Fig. 8. Time taken to converge to 20% of true $\sigma_{R_7}^a$.

second protocols to generate the correct diagnosis result, but the third protocol generates correct results without a coordinator. All three protocols, under certain assumptions, produce the same results as a centralized diagnoser. Our approach is similar to the third protocol, but, unlike the approach presented by Pencole and Cordier [2005], each individual local diagnoser needs to communicate only the minimal number of measurements, and not diagnosis results, from other diagnosers to generate globally correct diagnosis results.

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 in 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, such as in 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. In our work, we adopt the “shrinkage” approach presented in Liu and West [2000] to address this issue.

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.

Because the factors are conditionally independent, unlike distributed decentralized extended Kalman filters

(DDEKF) (see Mutambara [1998]), the failure of one distributed observer will not affect the estimations of other observers. Structural observability of each generated DBN-F guarantees that the distributed observers correctly estimate system behavior during nominal operation. However, structural observability does not guarantee that the system is observable with the fault magnitude introduced as an extra state variable.

6. DISCUSSION AND CONCLUSIONS

In this paper, we established how the distributed diagnosers truly generate globally correct results without any centralized coordinator, and through communicating the minimal number of measurements alone, and not individual diagnoses, unlike other previous work, such as Pencole and Cordier [2005]. The requirement for communicating partial diagnoses can be avoided because unlike other approaches, we have the knowledge of the global system model that is analyzed carefully for designing the diagnosers. However, there are several application domains, where the global models of large systems do not change, but they can greatly benefit from our distributed diagnosis scheme. Further, the DBN-Fs generated using our factoring scheme improves the efficiency of diagnosis without sacrificing accuracy of diagnosis.

In the future, we seek to investigate the important research problem of studying the observability of the faulty models once the extra fault variables are introduced. The problem of identifying the correct set of measurements such that the system is observable both during nominal and faulty operation, therefore, is an important research task. Next, we wish to apply our diagnosis approach to a large real-world system, to analyze the scalability and efficiency of our methodology. Finally, we would like to improve the efficiency of our diagnosis approach further by ensuring that the DBN-Fs are so chosen such that minimal number of fault hypotheses remain at the end of the Qual-FI.

REFERENCES

- X. Boyen and D. Koller. Tractable inference for complex stochastic processes. In *Proc. of the 14th Annual Conference on Uncertainty in Artificial Intelligence*, pages 33–42, 1998.
- 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.
- R. Debouk, S. Lafortune, and D. Teneketzis. Coordinated decentralized protocols for failure diagnosis of discrete event systems. *Discrete Event Dynamic System: Theory and Applications*, 10(1/2):33–86, January 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.
- 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.
- 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.
- J. Liu and M. West. Combined parameter and state estimation in simulation-based filtering. In J. F. G. De Freitas A. Doucet and N. J. Gordon, editors, *Sequential Monte Carlo Methods in Practice*. New York. Springer-Verlag, New York, 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.
- P. J. Mosterman and G. Biswas. Diagnosis of continuous valued systems in transient operating regions. *IEEE-SMCA*, 29(6):554–565, 1999.
- K. P. Murphy. *Dynamic Bayesian Networks: Representation, Inference, and Learning*. PhD thesis, University of California, Berkeley, 2002.
- A. Mutambara. *Decentralized Estimation and Control of Multisensor Systems*. CRC Press, 1998.
- 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.
- B. Ng and L. Peshkin. Factored particles for scalable monitoring. In *In Proceedings of the Eighteenth Conference on Uncertainty in Artificial Intelligence*, pages 370–377. Morgan Kaufmann, 2002.
- Yannick Pencole and Marie-Odile Cordier. A formal framework for the decentralised diagnosis of large scale discrete event systems and its application to telecommunication networks. *Artif. Intell.*, 164(1-2):121–170, 2005. ISSN 0004-3702.
- I. Roychoudhury, G. Biswas, and X. Koutsoukos. Comprehensive diagnosis of continuous systems using dynamic bayes nets. In *Proc. of the 19th International Workshop on Principles of Diagnosis*, pages 151–158, 2008.
- I. Roychoudhury, G. Biswas, and X. Koutsoukos. Designing distributed diagnosers for complex continuous systems. *IEEE Transactions on Automation Science and Engineering*, to appear, April 2009a.
- I. Roychoudhury, G. Biswas, and X. Koutsoukos. Efficient tracking for diagnosis using factored dynamic Bayesian networks. In *7th IFAC Symposium on Fault Detection, Supervision, and Safety of Technical Processes (SAFE-PROCESS 2009)*, to appear, 2009b.
- I. Roychoudhury, G. Biswas, and X. Koutsoukos. Factoring dynamic Bayesian networks based on structural observability. In *48th IEEE Conference on Decision and Control (CDC 2009)*, under review, 2009c.
- C. Seur and G. Dauphin-Tanguy. Bond graph approach for structural analysis of MIMO linear systems. *Journal of the Franklin Institute*, 328(1):55–70, 1991.
- V. Verma, G. Gordon, R. Simmons, and S. Thrun. Real-time fault diagnosis. *Robotics & Automation Magazine, IEEE*, 11(2):56–66, 2004.