

# Distributed Diagnosis in Uncertain Environments Using Dynamic Bayesian Networks

Indranil Roychoudhury, Gautam Biswas and Xenofon Koutsoukos  
ISIS, Dept. of EECS, Vanderbilt University, Nashville, TN 37235, USA  
{indranil.roychoudhury, gautam.biswas, xenofon.koutsoukos}@vanderbilt.edu

**Abstract**—Model-based diagnosis for industrial applications have to be efficient, and deal with modeling approximations and measurement noise. This paper presents a distributed diagnosis scheme, based on Dynamic Bayesian Networks (DBNs) that generates globally correct diagnosis results through local analysis, by only communicating a minimal number of measurements among diagnosers. We demonstrate experimentally that our distributed diagnosis scheme is computationally more efficient than its centralized counterpart, and it does not compromise the accuracy of the diagnosis results.

## I. INTRODUCTION

Online diagnosis schemes designed to ensure the safe and efficient operation of real-world engineering systems must be robust to uncertainties; efficient in their memory and computational requirements; scale well to changes in system configurations; and not suffer from single points of failure. Most centralized model-based diagnosis schemes suffer from some of these shortcomings, but distributed diagnosis schemes can address these drawbacks [1]–[3].

This paper presents a distributed scheme for diagnosing parametric faults in complex physical systems operating in uncertain environments using Dynamic Bayesian Networks (DBNs) [4]–[6]. Our distributed diagnosis scheme does not use a centralized coordinator, and each local diagnoser generates globally correct diagnosis results through local analysis, by only communicating a minimal number of measurements with other local diagnosers. The diagnoser design is based on factoring the bond graph (BG) model of the system into multiple independent, structurally observable, bond graph factors (BG-Fs) that are systematically converted into diagnosis models, i.e., DBN-Factors (DBN-Fs), used by the local diagnosers. Random variables in a DBN-F are guaranteed to be conditionally independent of the random variables in all other DBN-Fs given the chosen subset of communicated measurements considered as system inputs. We leverage this conditional independence among the DBN-F variables to derive separate particle filter (PF)-based inference algorithms [7] for fault detection, isolation, and identification. This quantitative diagnosis scheme is employed in combination with a qualitative fault isolation scheme to improve diagnosis efficiency.

In the remainder of this paper, we present our distributed diagnosis scheme, prove that our local diagnosers generate globally correct diagnosis results through local analysis without a centralized coordinator, and demonstrate experimentally using an electrical power system case study that our distributed

diagnosis scheme is computationally more efficient than its centralized counterpart without compromising the accuracy of the diagnosis results.

## II. MODELING FOR DIAGNOSIS

Our approach requires two diagnosis models: (1) *temporal causal graphs* (TCGs) for qualitative fault isolation [8], and (2) DBNs for fault detection and identification [6]. Both models are systematically derived from the system BG model [9].

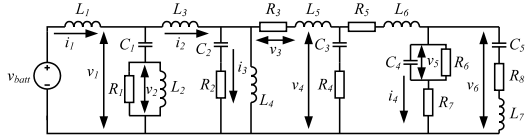
### A. Bond Graphs

BGs are parametric, topological models that capture energy exchange pathways in physical processes. The generic BG-elements 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 has an associated effort,  $e$ , and flow,  $f$ , variable, such that their product defines the power transferred through the bond. 0- and 1-junctions represent equal-effort and equal-flow connections, respectively. Fig. 1(b) shows the BG of a twelfth-order electrical circuit shown in Fig. 1(a). For every  $C$  element in integral causality, the corresponding state variable is the *displacement* variable,  $q$ , such that  $\dot{q} = f$ . Similarly for every  $I$  element in integral causality, the corresponding state variable is the *momentum* variable,  $p$ , such that  $\dot{p} = e$ .

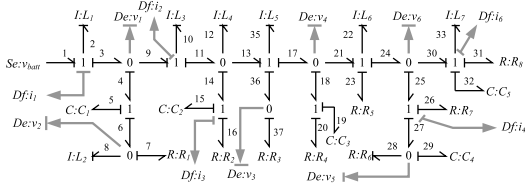
Faults are defined as changes in the nominal BG parameter values [6]. An *incipient fault* is a slow change in BG parameter,  $p$  (with nominal parameter value function,  $p(t)$ ), and modeled as  $p^{\pm i}(t) = p(t) \pm \Delta_p^i \cdot (t - t_f)$ ,  $t > t_f$ , where  $t_f$  is the time of fault occurrence,  $p^{\pm i}(t)$  is the temporal profile of parameter  $p$  with an incipient fault, and  $\Delta_p^i$  is a constant slope. An *abrupt fault* is modeled as an addition of a constant persistent bias term,  $\Delta_p^a \cdot p(t)$ , to the nominal parameter value,  $p(t)$ , i.e.,  $p^{\pm a}(t) = p(t) \pm \Delta_p^a \cdot p(t)$ ,  $t > t_f$ , where  $t_f$  is the time of fault occurrence,  $\Delta_p^a$  is the percentage change in the parameter expressed as a fraction, and  $p^{\pm a}(t)$  is the temporal profile of parameter  $p$  with an abrupt fault.

### B. Temporal Causal Graph

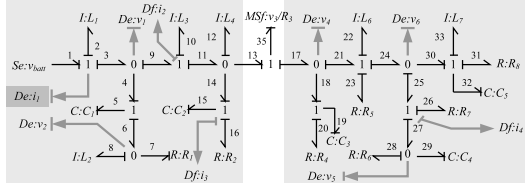
TCGs are graphs that capture the causal and temporal relations between system variables, through directed *edges* and their *labels*. The direction of a TCG edge and its label are based on *causality*, which establishes the cause and effect relationships between the  $e$  and  $f$  variables of a bond based on



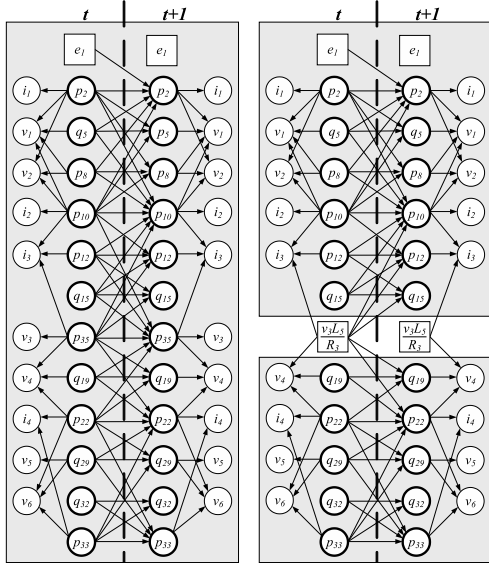
(a) Schematic.



(b) Bond graph.



(c) Two-Factored bond graph with imposed derivative causality.



(d) Dynamic Bayesian Network. (e) Two-Factored Dynamic Bayesian Network.

Fig. 1. Models of the twelfth-order electrical system.

constraints imposed by the incident BG elements. As shown in [8], algorithms that use the sequential causal assignment procedure (SCAP) to assign the causality in a BG [9] can be used to generate a TCG from a BG.

### C. Dynamic Bayesian Network

A DBN is defined as  $D = (\mathbf{X}, \mathbf{U}, \mathbf{Y})$ , where  $\mathbf{X}$ ,  $\mathbf{U}$ , and  $\mathbf{Y}$  are sets of stochastic random variables that denote (hidden) state variables, system input variables, and measured variables in the dynamic system, respectively [4]. 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 [5]. Fig. 1(d) 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.

## III. THE DISTRIBUTED DIAGNOSIS APPROACH

Decentralized diagnosis schemes can be broadly classified into three protocols presented in [2], where each local diagnoser is built from the global system model and uses only a subset of observable events. Our approach, similar to the third protocol, generates correct results without a coordinator. But, unlike the approach presented by [1], 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.

In our distributed diagnosis approach (Fig. 2), we factor a BG into structurally observable BG-Fs, and build a local diagnoser off the DBN-F and TCG-Factor (TCG-F) derived from each BG-F. Each local diagnoser performs three primary tasks [6]: (i) fault detection, (ii) qualitative fault isolation (Qual-FI), and (iii) quantitative fault hypothesis refinement and identification (Quant-FHRI).

A fault is *detected* when the residual, i.e., the difference between the observed (faulty) and estimated (nominal) values of a measurement, is determined to be statistically significant [10]. A PF scheme [7] implements the *nominal* observer for each DBN-F diagnoser. Fault detection triggers Qual-FI, which starts with a *hypotheses generation*, where all possible parameter changes that can explain the observed deviation are generated. The fault hypotheses are refined by comparing the *fault signatures* of the fault hypotheses, and removing from consideration, the fault hypotheses inconsistent with the observed deviations. Fault signatures are generated from the system TCG-F.

The Quant-FHRI scheme is invoked when 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. For each fault hypothesis that remains when Quant-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 [6]. Again, a PF scheme for each DBN-F fault model tracks the faulty observed behavior, taking as input the measurements from time  $t_d - \Delta_t^{max}$ , where  $\Delta_t^{max} \geq t_d - t_f$  is the maximum delay possible between the time of fault occurrence,  $t_f$ , and the time of fault detection,  $t_d$ . For each PF, 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. As more observations are obtained, ideally the PF using the correct fault model will eventually converge to the observed measurements, while the observations estimated using the incorrect fault models

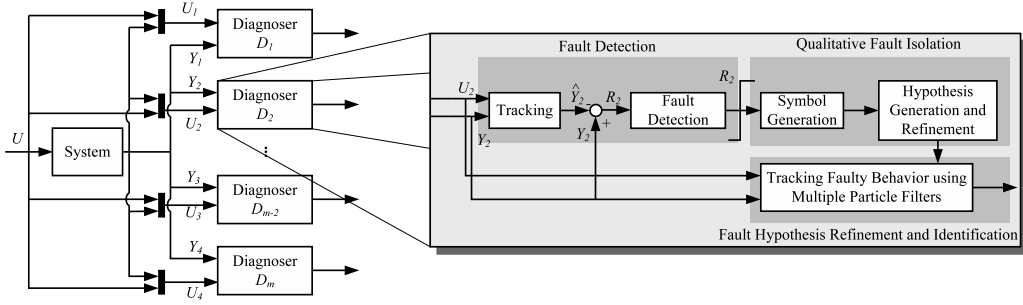


Fig. 2. The distributed diagnosis architecture.

would gradually deviate from the observed measurements. We 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,  $\Delta_p^a$ , or, the slope,  $\Delta_p^i$ ) is considered to be that estimated by the PF for the true fault model.

We start tracking faulty data from time  $t_d - \Delta_t^{\max}$  because once the fault has occurred, and till the magnitude of the fault is correctly identified, the system model is unknown. Tracking the observed measurements from before the occurrence of fault in the system is beneficial in terms of setting the initial state vector values. However, this implies that the process noise for each state variable must be set to a large enough value to avoid the “particle attrition” or “weight degeneracy” problem [11]. But having a large constant standard deviation would result in a large variance in estimated values. Fig. 3 illustrates our approach of using a standard deviation that varies linearly between  $\sigma_p^{\max}$  and  $\sigma_p^{\min}$  around the approximate time of fault occurrence. The heuristic for determining  $\sigma_p^{\max}$  and  $\sigma_p^{\min}$  are as follows: We assume a symmetric scale for fault magnitude increase and decreases, and assign  $\sigma_p^{\max} \geq \mathcal{P}/3$ , where  $\mathcal{P}$  is the nominal value of parameter  $p$  for a  $p^{-a}$  fault. The denominator 3 is based on normal distribution characteristics: 99.7% of the values in a normal distribution lies within three standard deviations of the mean. For a  $p^{-i}$  fault, we assign  $\sigma_p^{\min} \geq \mathcal{P}\Delta t/3$ , where  $\Delta t$  is the time difference between two slices of the DBN. For  $p^{+a}$  and  $p^{+i}$  faults, we assume the max possible values for  $\Delta_p^a$  and  $\Delta_p^i$  are denoted by  $\Delta_{p_{\max}}^a$  and  $\Delta_{p_{\max}}^i$ , respectively. So, for a  $p^{+a}$  fault,  $\sigma_p^{\max} \geq \Delta_{p_{\max}}^a/3$ , and for a  $p^{+i}$  fault,  $\sigma_p^{\max} \geq \Delta_{p_{\max}}^i \Delta t/3$ . Once  $\sigma_p^{\max}$  values have been determined, we usually set  $\sigma_p^{\min} = \sigma_p^{\max}/10$ . Note that in the incipient fault model, it is the noise that accounts for generating particles in the vicinity of the true parameter value at each time step. Hence, we must be careful to ensure that  $\sigma_p^{\min}$  is not smaller than the actual  $\Delta_p^i \cdot \Delta t$  for that parameter.

#### IV. DESIGNING THE LOCAL 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 first *factoring* the system BG into *maximal* number of

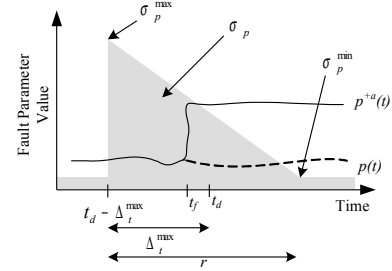


Fig. 3. Profile of standard for our particle filtering-based fault identification.

structurally observable [12] BG-Fs, and then generating a TCG-F and a DBN-F for each BG-F. The detailed procedure for generating the local diagnosers appears in [13].

The factoring approach identifies the system state variables that can be computed as algebraic functions of a subset of sensor measurements, and replaces these state variables by the algebraic functions. As the values of these state variables can now be computed at every time step, the dependence of across time relations between some of the state variables are removed, which decouples the computational model for each factor given the sensor measurements. Therefore, the inference algorithms for each factor are independent of the algorithms for other factors, facilitating our distributed diagnosis scheme.

Factoring BGs, therefore, corresponds to replacing energy-storage  $I$  and  $C$  elements by modulated sources of flow or effort, respectively, with the modulation factors computed as algebraic functions of observed measurements. Fig. 1(b) shows complete BG of the electrical circuit. The available measurements in this circuit are current values,  $i_1, i_2, \dots, i_4$  and voltages  $v_1, v_2, \dots, v_6$ . Fig. 1(c) shows the two BG-Fs that the global BG is factored into. The current through the inductor  $L_5$  is equal to  $v_3/R_3$ . Hence, we can replace  $f_{35}$  in Fig. 1(b) with modulated  $MS_f = v_3/R_3$  creating the two independent factors. If one or more of the resultant BG-Fs are structurally unobservable, they are *merged* with other BG-Fs till all of the BG-Fs are structurally observable.

In the maximal number of structurally observable BG-F configuration, each BG-F is converted into a DBN-F. By construction, the random variables in each DBN-F are conditionally independent from those in other DBN-Fs given the subset of

measurements now considered as system inputs. Because our DBN-Fs are generated from structurally observable BG-Fs, our factored inference scheme generates accurate inference results. Fig. 1(e) shows the DBN-Fs corresponding to the two BG-Fs in Fig. 1(c). Note that in the DBN-Fs, the state variable  $p_{35}$  is replaced by the input  $v_3L_5/R_3$ . Since,  $v_3/R_3$  can be *measured* at every time step, all causal links directed into  $p_{35}$  are removed. As a result, given  $v_3L_5/R_3$ , every variable in one DBN-F is conditionally independent of the variables in the other DBN-F. Thus, the two generated DBN-Fs are conditionally independent. In the factored DBN, we do not replace state variables, such as,  $p_2$  with  $i_1L_1$ , since this replacement does not yield any additional factors in Fig. 1(e). Moreover, we do not replace state variables  $p_{10}$  and  $q_{29}$  with  $i_2L_3$  and  $v_5C_4$ , respectively, since we assume that inductor  $L_3$ , and capacitor  $C_4$  can become faulty. We can see that the DBN-Fs shown in Fig. 1(e) map to the BG-Fs shown in Fig. 1(c).

Once we generate  $m$  DBN-Fs,  $D_1, D_2, \dots, D_m$ , from  $m$  structurally-observable BG-Fs,  $B_1, B_2, \dots, B_m$ , a local diagnoser,  $D_i$ , is constructed based off the DBN-F  $D_i$  and the TCG-F derived from each BG-F,  $B_i$ . Our distributed diagnosis approach (presented in the previous section) can be implemented by each  $D_i$ , independent of the other diagnosers.

## V. IMPLEMENTING THE LOCAL DIAGNOSERS

Each local diagnoser,  $D_i$ , receives the input signals  $\mathbf{U}_i$ , and the observed measurements  $\mathbf{Y}_i$  from the system since it is based off DBN-F,  $D_i$ . 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.,  $\mathbf{U}_i \cap \mathbf{Y} \neq \emptyset$ . Each  $D_i$  implements an independent PF-based observer on nominal DBN-F  $D_i$  for fault detection in Qual-FI; uses the TCG-F for hypothesis generation and refinement; and PFs applied to faulty DBN-Fs generated by extending  $D_i$  with faulty parameters as additional state variables for tracking faulty system behavior in Qual-FHRI.

Each of these PFs takes as inputs,  $\mathbf{U}_i$ , and estimates  $\mathbf{X}_i$  based on  $\mathbf{Y}_i$ . Only measurements  $(\cup_i \mathbf{U}_i) - \mathbf{U}$  are shared between the PF-based observers for each  $D_i$ . Further, the PF for the DBN-F  $D_i$  is designed to use  $a \frac{|\mathbf{X}_i|}{|\mathbf{X}|}$  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 PF with the maximum number of particles. The reduction of complexity is based on the assumptions that the sensors associated with measurements used to modulate the sources of energy in BG-Fs will not be faulty, and the components whose parameters are used in the algebraic functions are assumed not to fail. Therefore, there is a trade-off for robustness to gain efficiency.

Our local diagnosers are guaranteed to generate globally correct results through local analysis, without a centralized coordinator. By construction, a fault,  $\phi \in F_j$ , is only detected

TABLE I  
FAULT SIGNATURES FOR LOCAL DIAGNOSERS  $D_1$  AND  $D_2$

Diagnoser $D_1$						Diagnoser $D_2$					
Fault	$i_1$	$i_2$	$i_3$	$v_1$	$v_2$	Fault	$i_4$	$v_4$	$v_5$	$v_6$	
$C_2^{-a}, C_2^{-i}, R_2^{+a}, R_2^{+i}$	0	-	0	-	0+	0+					
$L_2^{-a}$	0+	0	-	0	-	$C_3^{-a}, R_4^{+a}$	0+	+	-	0+	
$L_2^{-i}$	0+	0	-	0	-	$C_3^{-i}, R_4^{+i}$	0+	0+	0+	0+	
$L_3^{-a}$	0+	+	+	-	-	$C_4^{-a}$	0-	0+	+	+	
$L_3^{-i}$	0+	+	+	-	-	$C_4^{-i}, R_6^{+a}, R_6^{+i}$	0-	0+	0+	0+	
$L_4^{-a}$	0+	0+	0+	0	-	$L_7^{-a}$	-	+	0	-*	
$L_4^{-i}$	0+	0+	+	0	-	$L_7^{-i}$	0-	0	-	0-	
$L_4^{+a}$	0+	0+	-	0	-	$R_7^{+a}$	0-	0+	0+	-	
$L_4^{+i}$	0+	0+	0	-	-	$R_7^{+i}$	0-	0+	0+	-	

by diagnoser  $D_j$ . All other diagnosers,  $D_k, k \neq j$ , will not detect the fault hence, they are not activated. In general, say the observer in diagnoser  $D_i$  uses the state space equations  $\hat{\mathbf{X}}_{i,t+1} = G_i(\mathbf{X}_i, \mathbf{U}_i)$ , and  $\hat{\mathbf{Y}}_i = H_i(\mathbf{X}_i, \mathbf{U}_i)$ . Two diagnosers  $D_j, D_k$  *communicate* a measurement  $Y \in \mathbf{Y}$  if  $Y \in \mathbf{U}_j \wedge Y \in \mathbf{U}_k$ , i.e., measurement  $Y$  is an input to both  $D_j$  and  $D_k$ . A fault in BG-F,  $B_k$  implies that functions  $G_k$  and  $H_k$  do not correctly represent the actual system any more. As a result,  $\hat{\mathbf{Y}}_k \not\approx \mathbf{Y}_k$ , 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 their shared inputs,  $(\mathbf{U}_j \cap \mathbf{U}_k) - \mathbf{U}$ , iff  $B_k$  and  $B_j$  communicate at least one measurement, but, since we adopt the single-fault assumption, and since by construction, two BG-Fs can never share any parameters, 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, the diagnoser does not get activated unless a fault is detected.

## VI. EXPERIMENTAL RESULTS

This section presents experimental results of applying our distributed diagnosis scheme on the electrical system shown in Fig. 1(a). Two local diagnosers,  $D_1$  and  $D_2$  are designed for this electrical circuit, for the top and bottom DBN-Fs shown in Fig. 1(e). The two diagnosers communicate voltage measurement  $v_3$  between each other. Table I shows the possible faults that must be diagnosed by each of the two diagnosers, and the fault signatures for each fault, given the measurements available to each diagnoser.

We present an experimental run for diagnosing an abrupt fault in  $C_2, C_2^{-a}$ , with  $\Delta_{C_2}^a = -0.9$ , introduced at time,  $t = 100$  s. A negative deviation is noticed in measurement  $i_3$  at  $t = 101.4$  s, which result in the fault hypotheses set,  $\{C_2^{-i}, C_2^{-a}, R_2^{+a}, R_2^{+i}, L_2^{-i}, L_2^{-a}, L_4^{-i}, L_4^{-a}\}$ . Subsequent changes in  $i_3$  and  $v_2$  the fault hypotheses are refined to  $\{C_2^{-i}, C_2^{-a}, R_2^{+a}, R_2^{+i}\}$  using the fault signatures in Table I. Qual-FI can produce no further refinements so Quant-FHRI is initiated. As shown in Fig. 4, the second diagnoser does not detect any fault. We start tracking the observed measurements from time  $t = 97.5$  s, and instantiate two PFs, one using a DBN-F model for fault  $C_2^{-i}/C_2^{-a}$ , and the other using a DBN-F model for fault  $R_2^{+i}/R_2^{+a}$ , with parameters  $C_2$  and  $R_2$  introduced as additional state variables in the nominal system DBN-Fs

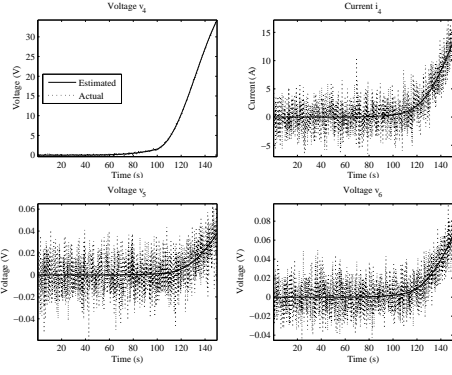
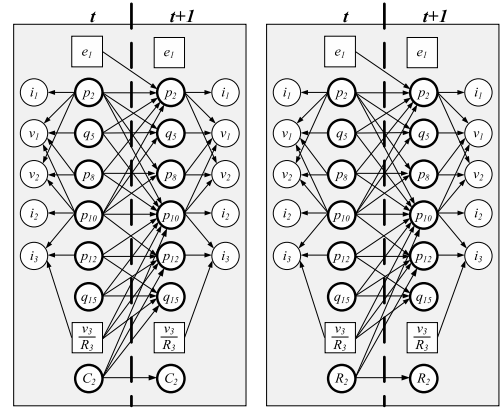


Fig. 4. Tracking observations in the presence of  $C_2^{-a}$  fault by diagnoser  $D_2$ .

shown in Fig. 1(e). Figs. 5(a) and 5(b) show the DBN-F fault models for  $C_2^{-i}/C_2^{-a}$  and  $R_2^{+i}/R_2^{+a}$ , respectively.

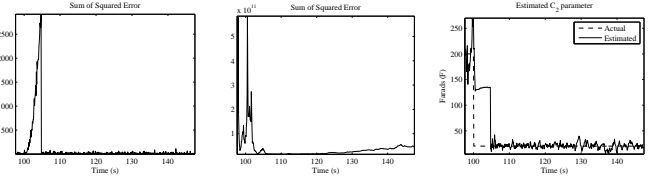
Fig. 6 shows the sum of mean squared estimation errors obtained using the two different fault models. For abrupt faults, at the time the fault is introduced, the abrupt fault parameter value is unknown. Hence, it takes some time for the correct fault model estimates to converge to the observed measurements, as we can see in the error plots obtained by the first fault model. A statistical test is employed on the sum of estimation errors across all the measurements to detect statistically significant sum of mean squared estimation errors. This statistical test detects a statistically significant sum of mean squared estimation error obtained by both the DBN-F fault models at times  $t = 101.3$  s and  $t = 101.9$  s, respectively. However, the sum of mean squared estimation errors from the  $R_2$  DBN fault model do not converge even after  $s_d = 150$  s, whereas, the sum of mean squared estimation errors from the  $C_2$  DBN fault model converges to the observed measurements from  $t = 105.0$  s. Hence the true fault is isolated to be  $C_2^{\pm a}/C_2^{\pm i}$  fault at  $t = 251.3$  s. In order to determine whether the fault is an abrupt or incipient fault in  $C_2$ , we run a window-based Z-test on the difference between the known nominal parameter value and the estimated state variable. At  $t = 171.7$  s, the statistical test shows that the estimated parameter evolves in a  $(-0)$  manner, implying it is an abrupt fault, and that it converges. By taking a mean of the values for 20 time steps after the abrupt fault is isolated, we obtain  $\Delta_{C_2}^a = -0.897$ . The actual value of  $\Delta_{C_2}^a$  is  $-0.900$ . Thus, there is a 0.33% error in estimating  $\Delta_{C_2}^a$ . The estimate for the faulty parameter is shown in Fig. 6(c). Notice the initial set of fault hypotheses generated in the distributed scheme is smaller than that generated in the centralized approach for the same fault experiment.

Table II summarizes the results of different distributed and centralized diagnosis experiments we ran on the electrical circuit example. For each experiment, we conducted 5 runs, and took the average of the time to fault detection, time to single fault isolation, time for the estimated parameter value to convergence to the true value, and the percentage error in the estimates of the true fault parameter. On comparing the results obtained from the centralized and distributed Bayesian



(a) DBN-F Fault model for  $C_2^{-a}/C_2^{-i}$ . (b) DBN-F Fault model for  $R_2^{+a}/R_2^{+i}$ .

Fig. 5. DBN-F Fault models for distributed diagnosis experiments.



(a) Sum of mean squared estimation errors obtained by  $C_2$  DBN fault models. (b) Sum of mean squared estimation errors obtained by  $R_2$  DBN fault models. (c) Estimate of  $C_2$  parameter obtained using the  $C_2$  DBN-F fault model.

Fig. 6. Quant-FHRI using  $C_2$  and  $R_2$  DBN-F fault models.

diagnosis experiments, we observe that the computational expense of the local diagnosers is less than that of the centralized diagnosers, since if each local diagnoser is implemented on a separate process, the worst case efficiency of our distributed diagnosis scheme is determined by the largest DBN-F fault model used for tracking faulty measurements, and, by construction, the largest DBN-F will still be smaller than the global DBN.

Also, compared to the centralized diagnosis approach, the distributed diagnosis approach results in comparable parameter estimation errors to the centralized diagnosis approach. In addition, the parameter estimates made by the distributed approach took longer to converge in terms of the number of measurement points required than the centralized scheme. We attribute this difference to the proportional distribution of particles based on the size of each factor, keeping the sum total of particles used by all the PFs the same. Moreover, the use of a noisy sensor to compute the value of a state variable also contributed to this degraded accuracy. The centralized diagnosis scheme has access to more sensors and the state-estimates are not as noisy as those computed in terms of measurements in the distributed scheme. If the individual local diagnosers are executed on different processors, then we can increase the number of particles for each diagnoser, and our intuition is that this will improve the estimation accuracy and identification time of the local diagnosers. Thus,

TABLE II  
RESULTS OF DIAGNOSIS EXPERIMENTS ON THE TWELFTH-ORDER  
ELECTRICAL CIRCUIT

Distributed Diagnosis with Particles Used Proportional to the Total Number of States Per Factor					
Fault	Magnitude	Detection Time (s)	Isolation Time (s)	Conv. Time (s)	% Mean Param. Est. Error
$C_2^{-a}$	-0.90	1.04	55.06	5.88	0.64
$L_3^{-a}$	-0.90	0.50	4.32	6.56	1.11
$C_3^{-a}$	-0.90	0.20	3.02	3.64	0.13
$R_7^{+a}$	+5.00	118.30	163.30	128.64	0.66
Centralized Diagnosis					
Fault	Magnitude	Detection Time (s)	Isolation Time (s)	Conv. Time (s)	% Mean Param. Est. Error
$C_2^{-a}$	-0.90	1.26	53.80	4.76	0.27
$L_3^{-a}$	-0.90	0.50	3.98	5.08	0.49
$C_3^{-a}$	-0.90	0.2	2.8	3.26	0.12
$R_7^{+a}$	+5.00	196.8	377.4	115.6	0.48

our experimental results on the twelfth-order electrical circuit illustrates the accuracy versus efficiency trade-off due to the factoring of the DBN into DBN-Fs.

## VII. DISCUSSION AND CONCLUSIONS

PFs have been used extensively for system health monitoring and diagnosis of hybrid systems [5], [14]. These approaches, however, do not alleviate the *sample impoverishment* problem, where low probability particles representing the faulty state are dropped during the re-sampling process. Several solutions have been proposed, e.g., [15] rank fault hypothesis based on their likelihoods, and report the most likely fault mode. Our approach to increasing them decreasing the variance of the unknown faulty parameters, similar to [11] addresses this issue successfully.

[16] propose an approach for combining look-ahead Rao-Blackwellised PFs (RBPFs) with Livingstone 3 (L3) for diagnosing faults in hybrid systems. 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.

The BK algorithm, presented in [17], creates the individual factors by eliminating causal links between weakly interacting subsystems. The belief state derived from the individual factors is an approximation of the true belief state, but the error is bounded. But the bounds may not be sufficiently precise for online diagnosis, leading to missed alarms and less precise diagnoses, or even false alarms and wrong diagnoses. The Factored Particle Filtering (FPF) scheme [18] reduces estimation errors by applying the particle filtering scheme to the BK factored inference approach. Our distributed estimation approach uses the particle filtering scheme for inference using DBNs and preserves the overall system dynamics in the factored form, and *does not* approximate the belief state. Hence, we produce accurate state estimates efficiently.

The effectiveness of our approach relies on the assumptions that the sensors associated with measurements converted to inputs are not faulty, and components whose parameters are used in the algebraic functions do not fail. In the future, we seek to relax these assumptions. In addition, we would like to analyze the accuracy, scalability, and efficiency of our methodology in large, industrial applications.

## REFERENCES

- [1] Y. Pencole and M.-O. Cordier, "A formal framework for the decentralised diagnosis of large scale discrete event systems and its application to telecommunication networks," *Artificial Intelligence*, vol. 164, no. 1-2, pp. 121-170, 2005.
- [2] R. Debouk, S. Lafortune, and D. Teneketzis, "Coordinated decentralized protocols for failure diagnosis of discrete event systems," *Discrete Event Dynamic System: Theory and Applications*, vol. 10, no. 1/2, pp. 33-86, January 2000.
- [3] I. Roychoudhury, G. Biswas, and X. Koutsoukos, "Designing distributed diagnosers for complex continuous systems," *IEEE Trans. on Automation Science and Engg.*, vol. 6, no. 2, pp. 277-290, April 2009.
- [4] K. P. Murphy, "Dynamic Bayesian networks: Representation, inference, and learning," Ph.D. dissertation, University of California, Berkeley, 2002.
- [5] U. Lerner, R. Parr, D. Koller, and G. Biswas, "Bayesian fault detection and diagnosis in dynamic systems," in *Proc. 17th AAAI Conf.*, 2000, pp. 531-537.
- [6] I. Roychoudhury, G. Biswas, and X. Koutsoukos, "Efficient tracking for diagnosis using factored dynamic Bayesian networks," in *Proc. 7th IFAC Symposium SAFEPROCESS 2009*.
- [7] D. Koller and U. Lerner, "Sampling in factored dynamic systems," in *Sequential Monte Carlo Methods in Practice*, A. Doucet, N. de Freitas, and N. Gordon, Eds. Springer-Verlag, 2001.
- [8] P. J. Mosterman and G. Biswas, "Diagnosis of continuous valued systems in transient operating regions," *IEEE-SMCA*, vol. 29, no. 6, pp. 554-565, 1999.
- [9] D. C. Karnopp, D. L. Margolis, and R. C. Rosenberg, *Systems Dynamics: Modeling and Simulation of Mechatronic Systems*, 3rd ed. NY: John Wiley, 2000.
- [10] E.-J. Manders, G. Biswas, P. J. Mosterman, L. A. Barford, and R. J. Barnett, "Signal interpretation for monitoring and diagnosis, a cooling system testbed," vol. 49, no. 3, pp. 503-509, 2000.
- [11] J. Liu and M. West, "Combined parameter and state estimation in simulation-based filtering," in *Sequential Monte Carlo Methods in Practice*. New York, A. Doucet, J. F. G. D. Freitas, and N. J. Gordon, Eds. Springer-Verlag, New York, 2000.
- [12] C. Sueur and G. Dauphin-Tanguy, "Bond graph approach for structural analysis of MIMO linear systems," *Journal of the Franklin Institute*, vol. 328, no. 1, pp. 55-70, 1991.
- [13] 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)*, to appear, 2009.
- [14] R. Dearden and D. Clancy, "Particle filters for real-time fault detection in planetary rovers," in *Proc. 12th Intl. Wkshp. on Principles of Diagnosis*, 2001, pp. 1-6.
- [15] V. Verma, G. Gordon, R. Simmons, and S. Thrun, "Real-time fault diagnosis," *Robotics & Automation Magazine, IEEE*, vol. 11, no. 2, pp. 56-66, 2004.
- [16] S. Narasimhan, R. Dearden, and E. Benazera, "Combining particle filters and consistency-based approaches for monitoring and diagnosis of stochastic hybrid systems," in *Proceedings of the 15th International Workshop on Principles of Diagnosis*, 2004.
- [17] X. Boyen and D. Koller, "Tractable inference for complex stochastic processes," in *Proc. 14th Conf. on Uncertainty in Artificial Intelligence*, 1998, pp. 33-42.
- [18] B. Ng and L. Peshkin, "Factored particles for scalable monitoring," in *Proceedings of the 18th Conference on Uncertainty in Artificial Intelligence*. Morgan Kaufmann, 2002, pp. 370-377.