Combating Curse of Dimensionality in Resilient Plant Monitoring Systems: Overlapping Decomposition with Intra- and Inter-Subplant Inferences

Humberto E. Garcia*, Semyon M. Meerkov† and Maruthi T. Ravichandran†

Abstract

This paper is intended to design and evaluate the performance of a power plant monitoring system (sensor network) that degrades gracefully under a cyber-physical attack. Such a system is referred to as resilient. While several features of the resilient system considered here have been developed in the previous work, the current paper addresses the "curse of dimensionality", which arises due to the exponential growth of the systems' state space as a function of the number of sensors. To combat this problem, we develop an overlapping decomposition of the power plant into a set of subplants, inducing a corresponding decomposition of the sensor network. Using both intra- and inter-subplant inferences, the paper mitigates the effect of information losses due to the decomposition. The efficiency of the resulting monitoring system is demonstrated by simulations.

I. Introduction

Resilient plant monitoring systems are sensor networks that degrade gracefully under cyber-physical attacks, which cause them to project misleading information. In the previous work [1]–[3], we have developed an approach to designing such systems based on the ideas of active data quality assessment [2] and rational control [4]–[6]. The former is used to identify sensors under cyber-physical attack, and the latter to adapt to the network state resulting in minimal entropy of the monitored plant assessment. While the steady state performance of such systems has been shown to be satisfactory, the transients have not: the adaptation time grows exponentially as a function of the network size (i.e., number of sensors in the network). This phenomenon, which arises in many engineering problems and which R. Bellman called the "curse of dimensionality", is the main topic addressed the current paper. Specifically, we develop an approach to combating the curse of dimensionality in resilient monitoring systems based

*Idaho National Laboratory, P.O. Box 1625, Idaho Falls, ID 83415-3675, USA. Email: Humberto.Garcia@inl.gov

†Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109, USA. Email: {smm, marutrav}@umich.edu

on the ideas of overlapping plant decomposition (developed in control-theoretic literature [7], [8]) and Dempster-Shafer combination rule [9] (used in artificial intelligence). The former induces an overlapping decomposition of the sensor network and, thus, reduces the state space of the adaptation; the latter allows for recovering information losses due to the decomposition.

To support this approach, we develop a plant monitoring system consisting of five layers: one of them remains the same as in the previous work (data quality assessment layer [2]), three others are modifications of those used previously (process variable assessment, adaptation, and plant assessment layers [1], [3]), and one is novel (knowledge fusion layer).

A specific system, in terms of which the development is carried out, is a simplified power plant shown in Figure 1, where B is the boiler, HT and LT are the high and low pressure turbines, respectively,

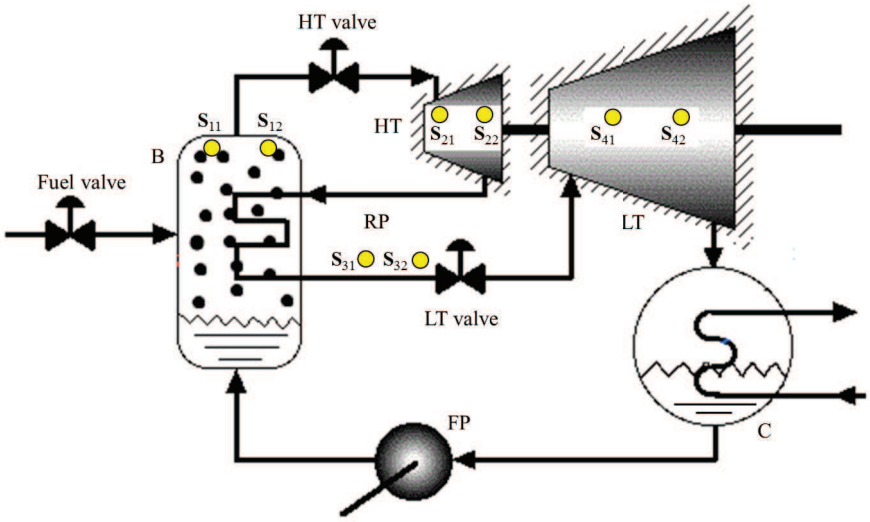


Fig. 1: Schematics of the power plant

RP is the reheat pipe, C is the condenser, FP is the feedwater pump, and $\mathbf{S}_{ij}$'s are the sensors; C and FP are assumed to operate normally, and, therefore, their sensors are not included; B, HT, RP, LT, C, and FP are referred to throughout this paper as *components* of the power plant. Having eight sensors, and assuming that the measurements of at least one of them is utilized for plant condition assessment, the cardinality of the sensor network state space is $2^8 - 1 = 255$. Decoupling the network into three overlapping subnetworks with four sensors each, leads to the cardinality $2^4 - 1 = 15$ and, thus, to dramatic improvement of the adaptation speed, in some cases without, as it is shown in this paper, loss in quality of plant condition assessment. Quantified by the Kullback-Leibler divergence [10], in all scenarios

considered, the resulting system exhibits high level of resiliency in comparison with non-resilient ones.

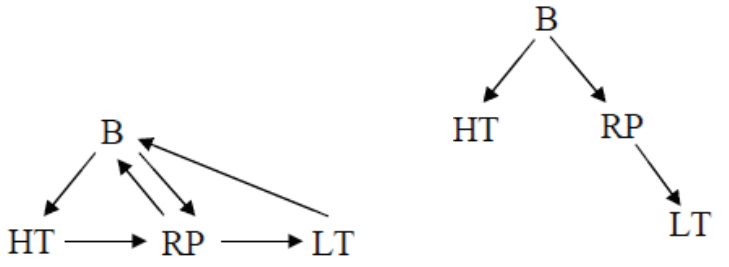
A brief review of the literature on other approaches to resilient plant monitoring systems is in order. (Humberto). However, the issue of combating the curse of dimensionality has not been addressed. This is carried out in this paper.

The remainder of this paper is structured as follows: The above mentioned overlapping decomposition architecture is developed in Section II. The model of the power plant considered is presented in Section III. The layers of the monitoring system are described in Sections IV-VIII. Results from numerical performance analysis of the overall system are presented in Section IX. Conclusions and directions of future work are given in Section X. The proofs are included in the Appendix.

II. Overlapping Decomposition Architecture

In this section, we describe the overlapping decomposition using as an example the power plant of Figure 1. Other plants could be addressed in a similar manner as well. Also, for the sake of simplicity, we assume that each component of the plant, which can be attacked, i.e., B, HT, RP, and LT, is characterized by a single process variable, e.g., its temperature, denoted as $\mathbf{V}_1$, $\mathbf{V}_2$, $\mathbf{V}_3$, and $\mathbf{V}_4$, respectively.

Mutual influences of the temperature among plant components can be represented by a *cyclic graph* shown in Figure 2(a). Assuming that B heat-generating capacity is large enough to maintain RP temperature independent of HT condition (normal or anomalous), the influence HT $\rightarrow$ RP can be ignored. Similarly,



(a) Cyclic graph representation (b) Tree graph representation

Fig. 2: Influence diagrams

under the above assumption, one can also ignore RP $\rightarrow$ B, since B is capable of maintaining its own temperature independent of HT and RP conditions. Further, if C heat-absorbing capacity is large enough to maintain constant water temperature at its outlet independent of LT conditions, the influence LT $\rightarrow$ B can also be ignored. Under these assumptions, the cyclic graph of Figure 2(a) is reduced to the *tree*

graph illustrated in Figure 2(b). This implies that all components of the power plant can be viewed as "serially connected", and the overlapping decomposition into three *subplants*, $\mathbf{G}_{\mathrm{I}}$, $\mathbf{G}_{\mathrm{II}}$, and $\mathbf{G}_{\mathrm{III}}$, shown in Figure 3, is possible. Such a decomposition, while reducing the dimensionality of the original problem,

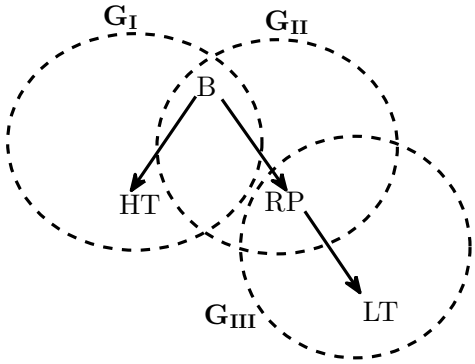


Fig. 3: Overlapping subplants $\mathbf{G}_{\mathrm{I}}$, $\mathbf{G}_{\mathrm{II}}$, and $\mathbf{G}_{\mathrm{III}}$

preserves all mutual influences of the process variables and leads (under appropriate conditions – see Subsection VII-B) to the same quality of plant condition assessment as the centralized architecture of Figure **??**. Note that overlapping decomposition, as a simplification tool, has been widely used in control theory for analysis and design of complex systems (see, for example, [7] and [8]).

As indicated in Figure 3, plant $\mathbf{G}$ is partitioned into three overlapping subplants, $\mathbf{G}_{\mathrm{I}}$, $\mathbf{G}_{\mathrm{II}}$, and $\mathbf{G}_{\mathrm{III}}$. This partitioning induces a corresponding partitioning of the sensor network $\mathbf{SN}$ into three *subnetworks*, $\mathbf{SN}_{\mathrm{I}}$, $\mathbf{SN}_{\mathrm{II}}$, and $\mathbf{SN}_{\mathrm{III}}$, consisting of $\{\mathbf{S}_{11}, \mathbf{S}_{12}, \mathbf{S}_{21}, \mathbf{S}_{22}\}$, $\{\mathbf{S}_{11}, \mathbf{S}_{12}, \mathbf{S}_{31}, \mathbf{S}_{32}\}$, and $\{\mathbf{S}_{31}, \mathbf{S}_{32}, \mathbf{S}_{41}, \mathbf{S}_{42}\}$, respectively. If X_k, $k =$ I, II, III, denotes the state space of each subnetwork, then the number of states in each of them is $2^4 - 1 = 15$, rather than 255 as in X. In this situation, if the evaluation of each state takes, as before, 10sec, a report to the operator is produced in less than 3min (rather than 42min, as in the centralized case). Note that under this decomposition, the aforementioned report would consist of four pmf's, $\hat{P}(\mathrm{B})$, $\hat{P}(\mathrm{HT})$, $\hat{P}(\mathrm{RP})$, and $\hat{P}(\mathrm{LT})$, rather than of a single pmf $\hat{P}(G)$.

Turning now to the issue of computing these pmf's, we introduce the *five-layer architecture* shown in Figure 4, where the additional, new layer is intended to utilize *inter-subplant inferences* to produce assessments similar to those that could be produced within the centralized architecture of Figure **??**. We refer to this new layer as the *knowledge fusion layer*.

The system of Figure 4 consists of three parallel sub-architectures, where the first three layers in each of them are similar, but not identical, to DQ assessment, process variable assessment, and A/RC layers of the centralized architecture. While the first of these layers operates in the same manner as in the

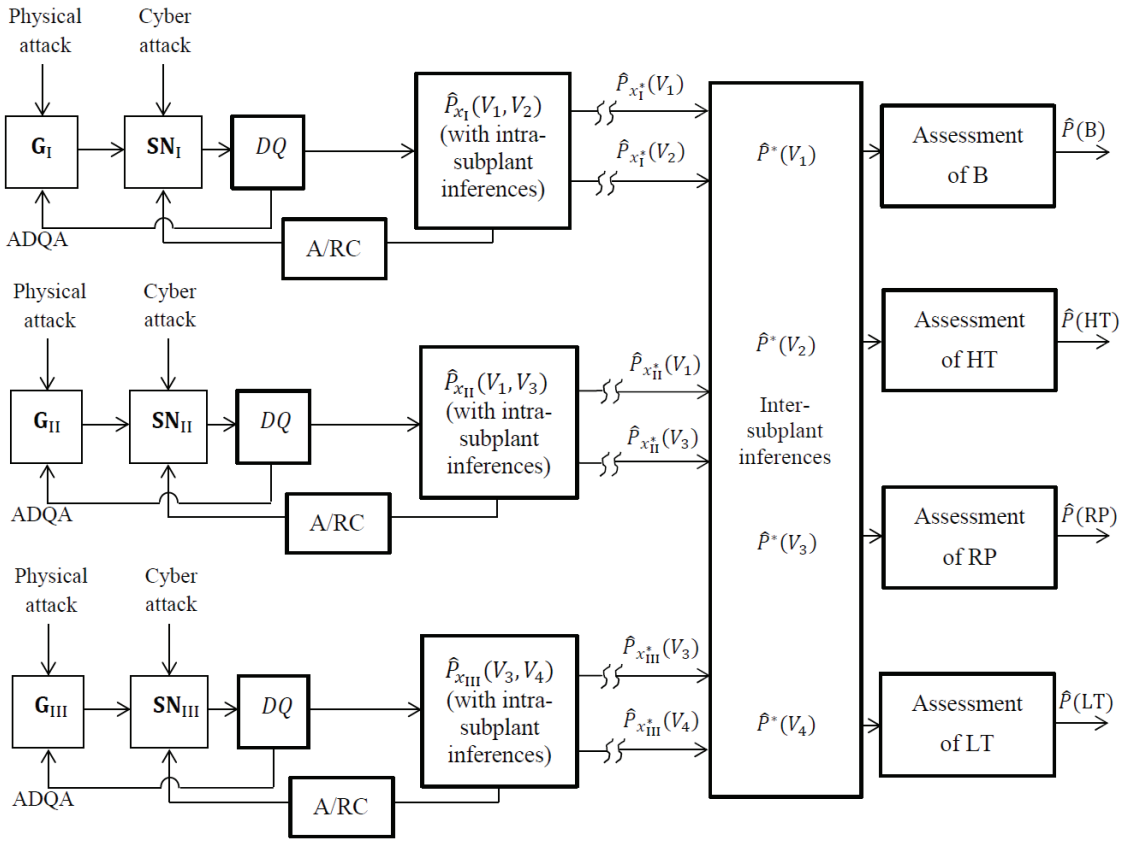


Fig. 4: Overlapping five-layer monitoring system architecture

centralized architecture, the latter two are different. Namely, the process variable assessment layer uses *intra-subplant* inferences to calculate joint pmf's of process variables of each subplant (i.e., $\hat{P}(V_1, V_2)$ in $\mathbf{G}_{\mathrm{I}}$, $\hat{P}(V_1, V_3)$ in $\mathbf{G}_{\mathrm{II}}$, and $\hat{P}(V_3, V_4)$ in $\mathbf{G}_{\mathrm{III}}$). The A/RC layer is different because it uses these joint pmf's as a basis of adaptation, rather than the plant pmf. The reason is that the subplant pmf's are not yet available at this stage of computation, since they require inter-subplant inferences. These inferences are utilized in the fourth (knowledge fusion) layer, which combines information from all subplants and evaluates pmf's of all process variables involved in the power plant. Finally, based on these pmf's and statistical models of power plant components, the fifth layer assesses their conditions and provides a report to the plant operator.

As mentioned before, details of these calculations are described in Sections IV-VIII.

III. Modeling

In this section, we model process variables of the power plant shown in Figure 1 for various conditions – normal and anomalous. While we assume that anomalies in each component of the plant occur independently, the process variables under various conditions are, in fact, coupled as shown in Figure 2(b), and this coupling leads to the models of process variables developed below.

A. Models of process variables

As mentioned before, states of B, HT, RP, and LT are assumed to be characterized by the following process variables: boiler temperature ($\mathbf{V}_1$), high pressure turbine temperature ($\mathbf{V}_2$), reheat pipe temperature ($\mathbf{V}_3$), and low pressure turbine temperature ($\mathbf{V}_4$), respectively. It should be noted that this assumption is not restrictive, and more complex models of power plants can be examined using the techniques developed in this paper.

Statistical models of the process variables and their relationships with the states of associated components are introduced as follows:

1) Model of boiler temperature: Let $\tilde{V}_1$ be a continuous random variable that characterizes values of $\mathbf{V}_1$. Let $f_{\tilde{V}_1}(\tilde{v}_1)$ denote the probability density function (pdf) of $\tilde{V}_1$ with the support

$$\tilde{V}_1 \in [V^{\mathrm{B}}_{\min}, V^{\mathrm{B}}_{\max}]. \tag{1}$$

In general, depending on the values taken by process variables, their states are often classified as normal or anomalous. Regarding B, low temperature values are assumed to occur due to a fracture in the insulation. Accordingly, the interval (1) is partitioned into the following two regions:

$$\mathrm{L_B}: \ [V^{\mathrm{B}}_{\min}, R^{\mathrm{B}}) \text{ and } \mathrm{N_B}: \ [R^{\mathrm{B}}, V^{\mathrm{B}}_{\max}], \tag{2}$$

where $\mathbf{V}_1$ is viewed as Low ($\mathrm{L_B}$) and Normal ($\mathrm{N_B}$), respectively. To characterize the state of $\mathbf{V}_1$ (according to (2)), we define the discrete random variable V_1 whose universal set is given by

$$\Sigma_1 \stackrel{\triangle}{=} \{\mathrm{L_B}, \mathrm{N_B}\}. \tag{3}$$

The probability mass function (pmf) of V_1 can be calculated using $f_{\tilde{V}_1}(\tilde{v}_1)$, $V^{\mathrm{B}}_{\min}$, $V^{\mathrm{B}}_{\max}$, and R^{B} as

$$\begin{aligned} P(V_1 = \mathrm{L_B}) &= \int_{V^{\mathrm{B}}_{\min}}^{R^{\mathrm{B}}} f_{\tilde{V}_1}(\tilde{v}_1)\mathrm{d}\tilde{v}_1, \\ P(V_1 = \mathrm{N_B}) &= \int_{R^{\mathrm{B}}}^{V^{\mathrm{B}}_{\max}} f_{\tilde{V}_1}(\tilde{v}_1)\mathrm{d}\tilde{v}_1. \end{aligned} \tag{4}$$

The d.c. gain of $\mathbf{V}_1$ with respect to the fuel valve position is assumed to be a continuous piecewise linear function of $\tilde{V}_1$, and is expressed as

$$\alpha_1(\tilde{V}_1) = \begin{cases} \alpha_{\mathrm{L_B}}(\tilde{V}_1), & \text{if } \tilde{V}_1 \in [V^{\mathrm{B}}_{\min}, R^{\mathrm{B}}) \\ \alpha_{\mathrm{N_B}}(\tilde{V}_1), & \text{if } \tilde{V}_1 \in [R^{\mathrm{B}}, V^{\mathrm{B}}_{\max}], \end{cases} \tag{5}$$

where $\alpha_{\mathrm{L_B}}(\tilde{V}_1)$ and $\alpha_{\mathrm{N_B}}(\tilde{V}_1)$ are linear in their respective domains and $\alpha_{\mathrm{L_B}}(R^{\mathrm{B}}) = \alpha_{\mathrm{N_B}}(R^{\mathrm{B}})$.

Thus, the model of $\mathbf{V}_1$ is defined by the pdf of $\tilde{V}_1$, pmf of V_1, and d.c. gain $\alpha_1(\tilde{V}_1)$, where $\tilde{V}_1 \in [V^{\mathrm{B}}_{\min}, V^{\mathrm{B}}_{\max}]$.

2) Model of high pressure turbine temperature: Let the random variable $\tilde{V}_2$ characterize the values of process variable $\mathbf{V}_2$. The pdf of $\tilde{V}_2$ is specified by $f_{\tilde{V}_2}(\tilde{v}_2)$ with the support

$$\tilde{V}_2 \in [V^{\mathrm{HT}}_{\min}, V^{\mathrm{HT}}_{\max}]. \tag{6}$$

Regarding the anomalous operation of HT, it is assumed to be similar to that of B, i.e., damage in the insulation. As before, let the discrete random variable V_2 characterize the state of $\mathbf{V}_2$. To define the state space of V_2, we also take into consideration, apart from the HT anomaly, the serial connection between B and HT (introduced in Section II), which implies that the condition of the former affects that of the latter. With the assumption that damage in B results in a larger drop in temperature than damage in HT alone, the universal set of V_2 is specified as

$$\Sigma_2 \triangleq \{\mathrm{VL_{HT}}, \mathrm{L_{(1)HT}}, \mathrm{L_{(2)HT}}, \mathrm{N_{HT}}\}, \tag{7}$$

where $\mathrm{VL_{HT}}$ ('Very Low') occurs when both HT and B are damaged; $\mathrm{L_{(1)HT}}$ when only B is damaged; $\mathrm{L_{(2)HT}}$ when only HT is damaged; and $\mathrm{N_{HT}}$ when neither B nor HT are damaged. The regions that specify the states of $\mathbf{V}_2$ are given by

$$\begin{aligned} &\mathrm{VL_{HT}}: \ [V^{\mathrm{HT}}_{\min}, R^{\mathrm{HT}}_1), \ \mathrm{L_{(1)HT}}: \ [R^{\mathrm{HT}}_1, R^{\mathrm{HT}}_2), \\ &\mathrm{L_{(2)HT}}: \ [R^{\mathrm{HT}}_2, R^{\mathrm{HT}}_3), \ \mathrm{N_{HT}}: \ [R^{\mathrm{HT}}_3, V^{\mathrm{HT}}_{\max}], \end{aligned} \tag{8}$$

where $R^{\mathrm{HT}}_1 < R^{\mathrm{HT}}_2 < R^{\mathrm{HT}}_3$. The pmf of V_2 can be calculated using $f_{\tilde{V}_2}(\tilde{v}_2)$, $V^{\mathrm{HT}}_{\min}$, $V^{\mathrm{HT}}_{\max}$, and R^{HT}_i, $i \in \{1, 2, 3\}$, as in (4).

The d.c. gain of $\mathbf{V}_2$ with respect to the HT valve position is assumed to be a continuous piecewise

linear function of $\tilde{V}_2$, and is expressed as

$$\alpha_2(\tilde{V}_2) = \begin{cases} \alpha_{\mathrm{VL_{HT}}}(\tilde{V}_2), & \text{if } \tilde{V}_2 \in [V_{\min}^{\mathrm{HT}}, R_1^{\mathrm{HT}}) \\ \alpha_{\mathrm{L_{(1)HT}}}(\tilde{V}_2), & \text{if } \tilde{V}_2 \in [R_1^{\mathrm{HT}}, R_2^{\mathrm{HT}}) \\ \alpha_{\mathrm{L_{(2)HT}}}(\tilde{V}_2), & \text{if } \tilde{V}_2 \in [R_2^{\mathrm{HT}}, R_3^{\mathrm{HT}}) \\ \alpha_{\mathrm{N_{HT}}}(\tilde{V}_2), & \text{if } \tilde{V}_2 \in [R_3^{\mathrm{HT}}, V_{\max}^{\mathrm{HT}}], \end{cases} \tag{9}$$

where $\alpha_{\mathrm{VL_{HT}}}(\tilde{V}_2)$, $\alpha_{\mathrm{L_{(1)HT}}}(\tilde{V}_2)$, $\alpha_{\mathrm{L_{(2)HT}}}(\tilde{V}_2)$, and $\alpha_{\mathrm{N_{HT}}}(\tilde{V}_2)$ are linear in their respective domains and $\alpha_{\mathrm{VL_{HT}}}(R_1^{\mathrm{HT}}) = \alpha_{\mathrm{L_{(1)HT}}}(R_1^{\mathrm{HT}})$, $\alpha_{\mathrm{L_{(1)HT}}}(R_2^{\mathrm{HT}}) = \alpha_{\mathrm{L_{(2)HT}}}(R_2^{\mathrm{HT}})$, and $\alpha_{\mathrm{L_{(2)HT}}}(R_3^{\mathrm{HT}}) = \alpha_{\mathrm{N_{HT}}}(R_3^{\mathrm{HT}})$.

Thus, the model of $\mathbf{V}_2$ is defined by the pdf of $\tilde{V}_2$, pmf of V_2, and d.c. gain $\alpha_2(\tilde{V}_2)$, where $\tilde{V}_2 \in [V_{\min}^{\mathrm{HT}}, V_{\max}^{\mathrm{HT}}]$.

3) Model of reheat pipe temperature: Anomalous operation of RP is, as before, assumed to occur due to damage in its insulation. Further, as mentioned in Section II, we assume that B and RP are serially connected while ignoring the effect of HT on the latter.

Let the continuous random variable $\tilde{V}_3$ characterize values of $\mathbf{V}_3$. The pdf of $\tilde{V}_3$ is denoted by $f_{\tilde{V}_3}(\tilde{v}_3)$ with the support

$$\tilde{V}_3 \in [V_{\min}^{\mathrm{RP}}, V_{\max}^{\mathrm{RP}}]. \tag{10}$$

As with previous statistical models, let the discrete random variable V_3 characterize the state of $\mathbf{V}_3$. The universal set of V_3 is specified as

$$\Sigma_3 \triangleq \{\mathrm{VL_{RP}}, \mathrm{L_{(1)RP}}, \mathrm{L_{(2)RP}}, \mathrm{N_{RP}}\}, \tag{11}$$

where $\mathrm{VL_{RP}}$ occurs when both B and RP are damaged; $\mathrm{L_{(1)RP}}$ when only B is damaged; $\mathrm{L_{(2)RP}}$ when only RP is damaged; and $\mathrm{N_{RP}}$ when neither B nor RP are damaged. The regions that represent the states of $\mathbf{V}_3$ are given by

$$\begin{aligned} &\mathrm{VL_{RP}}: [V_{\min}^{\mathrm{RP}}, R_1^{\mathrm{RP}}),\ \mathrm{L_{(1)RP}}: [R_1^{\mathrm{RP}}, R_2^{\mathrm{RP}}), \\ &\mathrm{L_{(2)RP}}: [R_2^{\mathrm{RP}}, R_3^{\mathrm{RP}}),\ \mathrm{N_{RP}}: [R_3^{\mathrm{RP}}, V_{\max}^{\mathrm{RP}}], \end{aligned} \tag{12}$$

where $R_1^{\mathrm{RP}} < R_2^{\mathrm{RP}} < R_3^{\mathrm{RP}}$. The pmf of V_3 can be calculated as before.

The d.c. gain of $\mathbf{V}_3$ with respect to the fuel valve position is assumed to be a continuous piecewise

linear function of $\tilde{V}_3$, and is expressed as

$$\alpha_3(\tilde{V}_3) = \begin{cases} \alpha_{\mathrm{VL_{RP}}}(\tilde{V}_3), & \text{if } \tilde{V}_3 \in [V_{\min}^{\mathrm{RP}}, R_1^{\mathrm{RP}}) \\ \alpha_{\mathrm{L_{(1)RP}}}(\tilde{V}_3), & \text{if } \tilde{V}_3 \in [R_1^{\mathrm{RP}}, R_2^{\mathrm{RP}}) \\ \alpha_{\mathrm{L_{(2)RP}}}(\tilde{V}_3), & \text{if } \tilde{V}_3 \in [R_2^{\mathrm{RP}}, R_3^{\mathrm{RP}}) \\ \alpha_{\mathrm{N_{RP}}}(\tilde{V}_3), & \text{if } \tilde{V}_3 \in [R_3^{\mathrm{RP}}, V_{\max}^{\mathrm{RP}}], \end{cases} \tag{13}$$

where $\alpha_{\mathrm{VL_{RP}}}(\tilde{V}_3)$, $\alpha_{\mathrm{L_{(1)RP}}}(\tilde{V}_3)$, $\alpha_{\mathrm{L_{(2)RP}}}(\tilde{V}_3)$, and $\alpha_{\mathrm{N_{RP}}}(\tilde{V}_3)$ are linear in their respective domains and $\alpha_{\mathrm{VL_{RP}}}(R_1^{\mathrm{RP}}) = \alpha_{\mathrm{L_{(1)RP}}}(R_1^{\mathrm{RP}})$, $\alpha_{\mathrm{L_{(1)RP}}}(R_2^{\mathrm{RP}}) = \alpha_{\mathrm{L_{(2)RP}}}(R_2^{\mathrm{RP}})$, and $\alpha_{\mathrm{L_{(2)RP}}}(R_3^{\mathrm{RP}}) = \alpha_{\mathrm{N_{RP}}}(R_3^{\mathrm{RP}})$.

Thus, the model of $\mathbf{V}_3$ is defined by the pdf of $\tilde{V}_3$, pmf of V_3, and d.c. gain $\alpha_3(\tilde{V}_3)$, where $\tilde{V}_3 \in [V_{\min}^{\mathrm{RP}}, V_{\max}^{\mathrm{RP}}]$.

4) Model of low pressure turbine temperature: Since LT operates at a low pressure, we assume that its insulation is immune to damage. Indeed, anomalous operation is considered to occur when LT transfers energy to the shaft with poor efficiency, i.e., heat is accumulated in the turbine, rather than being converted into mechanical energy. Consequently, under abnormal operation, LT temperature $\mathbf{V}_4$ is higher than normal. In addition, it is assumed that any rise in LT temperature is not sufficient to compensate for the loss of heat from either a damaged B, or a damaged RP, or both.

Let the continuous random variable $\tilde{V}_4$ characterize the values of $\mathbf{V}_4$. The pdf of $\tilde{V}_4$ is specified by $f_{\tilde{V}_4}(\tilde{v}_4)$ with the support

$$\tilde{V}_4 \in [V_{\min}^{\mathrm{LT}}, V_{\max}^{\mathrm{LT}}]. \tag{14}$$

Let the discrete random variable V_4 represent the state of $\mathbf{V}_4$. Taking into account the serial connection between B, RP, and LT (introduced in Section II), the universal set of V_4 is expressed as

$$\Sigma_4 \triangleq \{\mathrm{VL_{(1)LT}}, \mathrm{VL_{(2)LT}}, \mathrm{L_{(1)LT}}, \mathrm{L_{(2)LT}}, \mathrm{M_{(1)LT}}, \mathrm{M_{(2)LT}}, \mathrm{N_{LT}}, \mathrm{H_{LT}}\}, \tag{15}$$

where

$\mathrm{VL_{(1)LT}}$ implies LT is operating normally, RP is damaged, and B is damaged;

$\mathrm{VL_{(2)LT}}$ implies LT is malfunctioning, RP is damaged, and B is damaged;

$\mathrm{L_{(1)LT}}$ implies LT is operating normally, RP is operating normally, and B is damaged;

$\mathrm{L_{(2)LT}}$ implies LT is malfunctioning, RP is operating normally, and B is damaged;

$\mathrm{M_{(1)LT}}$ ('Medium') implies LT is operating normally, RP is damaged, and B is operating normally;

$\mathrm{M_{(2)LT}}$ implies LT is malfunctioning, RP is damaged, and B is operating normally;

$\mathrm{N_{LT}}$ implies LT is operating normally, RP is operating normally, and B is operating normally;

$\mathrm{H_{LT}}$ ('High') implies LT is malfunctioning, RP is operating normally, and B is operating normally.

The pmf of V_4 can be calculated like in (4).

As before, the d.c. gain of $\mathbf{V}_4$ with respect to the LT valve position is assumed to be a continuous piecewise linear function of $\tilde{V}_4$, and is denoted by $\alpha_4(\tilde{V}_4)$.

Thus, the model of $\mathbf{V}_4$ is defined by the pdf of $\tilde{V}_4$, pmf of V_4, and d.c. gain $\alpha_4(\tilde{V}_4)$, where $\tilde{V}_4 \in [V_{\min}^{\mathrm{LT}}, V_{\max}^{\mathrm{LT}}]$.

B. Modeling the mutual influences of process variables

The serial connections illustrated in Figure 2(b) are statistically modeled as conditional pmf's of the following pairs of process variables: $(\mathbf{V}_1, \mathbf{V}_2)$, $(\mathbf{V}_1, \mathbf{V}_3)$, and $(\mathbf{V}_3, \mathbf{V}_4)$. The conditional pmf $P(V_1|V_2)$, as it follows from Subsection III-A, is specified as

$$P(V_1|V_2) = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \tag{16}$$

where the rows and columns represent the states of V_1 and V_2, respectively. As it follows from Subsection III-A, if V_1 is $\mathrm{L_B}$, then V_2 can be either $\mathrm{VL_{HT}}$ or $\mathrm{L_{(1)HT}}$ with equal probabilities, and if V_1 is $\mathrm{N_B}$, then V_2 can be either $\mathrm{L_{(2)HT}}$ or $\mathrm{N_{HT}}$, once again, with equal probabilities. Therefore,

$$P(V_2|V_1) = \begin{bmatrix} 0.5 & 0 \\ 0.5 & 0 \\ 0 & 0.5 \\ 0 & 0.5 \end{bmatrix}. \tag{17}$$

Regarding $(\mathbf{V}_1, \mathbf{V}_3)$, as it follows from Subsection III-A, the conditional pmf's $P(V_1|V_3)$ and $P(V_3|V_1)$ are the same as (16) and (17), respectively. Regarding $(\mathbf{V}_3, \mathbf{V}_4)$, as it follows from Subsection III-A, the conditional pmf's $P(V_4|V_3)$ and $P(V_3|V_4)$ are

$$P(V_4|V_3) = \begin{bmatrix} 0.5 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 \\ 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 0.5 \end{bmatrix} \tag{18}$$

and

$$P(V_3|V_4) \;=\; \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}. \tag{19}$$

As mentioned in Section I, the conditional pmf's described above are utilized to obtain inferences about the state of one process variable, given the state of another.

C. Model of sensor subnetworks

As illustrated in Figure 1, each process variable $\mathbf{V}_i$, $i \in \{1,2,3,4\}$, is assumed to be monitored by two sensors, denoted by $\mathbf{S}_{ij}$, $j \in \{1,2\}$, where i and j index the process variable and sensor, respectively. Let the continuous random variable $\tilde{S}_{ij}$, whose pdf is $f_{\tilde{S}_{ij}}(\tilde{s}_{ij})$, characterize the measurements of sensor $\mathbf{S}_{ij}$. The range of $\tilde{S}_{ij}$ is identical to that of $\tilde{V}_i$ (see (1), (6), (10), and (14)). Let the discrete random variable S_{ij}, with universal set identical to that of V_i (see (3), (7), (11) and (15)), represent the state of $\mathbf{S}_{ij}$. The pmf of S_{ij} is calculated as that of V_i (see Subsection III-A).

Given that sensors may be attacked, we assign to each sensor a scalar value termed as data quality, $DQ \in [0,1]$, where, as before, $DQ = 1$ implies that the sensor measurements are completely trustworthy and $DQ = 0$ indicates that they are completely untrustworthy. While the method of assigning DQ's is described in Section IV, data quality is used here to specify a model of the sensors.

Since DQ is not a statistical quantity, a model of its effect on the relationship between random variables V_i and S_{ij} should be introduced. To accomplish this, we define the sensor *believability*, β_{ij}, as

$$\beta_{ij} = \frac{|\Sigma_i| - 1}{|\Sigma_i|} DQ_{ij} + \frac{1}{|\Sigma_i|},\ i \in \{1,2,3,4\},\ j \in \{1,2\}, \tag{20}$$

where $|\Sigma_i|$ represents cardinality of Σ_i, and sets Σ_i are given by (3), (7), (11), and (15). Equation (20) implies that if $DQ = 1$, believability is 1, while if $DQ = 0$, believability is $\frac{1}{|\Sigma_i|}$, implying that all states of the process variable $\mathbf{V}_i$ are equiprobable. The coupling between V_i and S_{ij} is postulated using believability as follows:

$$\begin{aligned} P\left(V_i = \sigma_i \,|S_{ij} = \sigma_i\right) &= \beta_{ij}, \\ P\left(V_i = \bar{\sigma}_i \,|S_{ij} = \sigma_i\right) &= \tfrac{1-\beta_{ij}}{|\Sigma_i|-1}, \end{aligned} \tag{21}$$

where $\bar{\sigma}_i$ implies 'not σ_i' and $\sigma_i, \bar{\sigma}_i \in \Sigma_i$, $i \in \{1,2,3,4\}$, $j \in \{1,2\}$. For example, consider sensor $\mathbf{S}_{11}$, which monitors $\mathbf{V}_1$. Since $|\Sigma_1| = 2$, believability is $\beta_{11} = \frac{DQ_{11}}{2} + \frac{1}{2}$, and, therefore, $P(V_1|S_{11})$ is

$$P(V_1|S_{11}) = \begin{bmatrix} \beta_{11} & 1-\beta_{11} \\ 1-\beta_{11} & \beta_{11} \end{bmatrix}, \tag{22}$$

where the rows and columns represent the states of $\mathbf{V}_1$ and $\mathbf{S}_{11}$, respectively. If $DQ_{11} = 1$, $P(V_1|S_{11}) = \mathrm{diag}(1,1)$, and if $DQ_{11} = 0$, both states of $\mathbf{V}_1$ are equally probable irrespective of the state of $\mathbf{S}_{11}$.

As it follows from Section I, the state space of sensor subnetwork $\mathbf{SN}_k$, $k \in \{\mathrm{I}, \mathrm{II}, \mathrm{III}\}$, is expressed as

$$X_k = \{(1000)_k, (0100)_k, ..., (1111)_k\}, \ k \in \{\mathrm{I}, \mathrm{II}, \mathrm{III}\}, \tag{23}$$

where the cardinality of X_k is $2^4 - 1 = 15$.

Thus, the model of sensor subnetworks is characterized by the model of individual sensors and the state space X_k, $k \in \{\mathrm{I}, \mathrm{II}, \mathrm{III}\}$, of each subnetwork.

D. Model of attacker

As far as the attacker is concerned, it is assumed that sensor measurements are modified in order to project misleading information. In formal terms, this implies that the attacker modifies $f_{\tilde{S}_{ij}}(\tilde{s}_{ij})$, $i \in \{1,2,3,4\}$, $j \in \{1,2\}$, by changing its variance, or expected value, or both. Our preliminary investigation indicated that modifying expected values is more damaging for resilient monitoring than modifying variances. Therefore, the model of the attacker considered in this paper is that for a sensor under attack,

$$E(\tilde{S}_{ij}) \neq E(\tilde{V}_i), \ i \in \{1,2,3,4\}, \ j \in \{1,2\}, \tag{24}$$

where $E(.)$ denotes the expected value. This implies, for example, that, while B temperature, $\mathbf{V}_1$, is in state $\mathrm{N_B}$, sensor $\mathbf{S}_{11}$ may project a signal indicating that $\mathbf{V}_1$ is in state $\mathrm{L_B}$.

The mean value-based attacker (24) is considered throughout this paper. We note, however, that other models of the attacker could be considered using the approach of this work.

E. Plant model

The plant model consists of models for B, HT, RP, and LT. Since, as it is mentioned at the beginning of this section, the anomalies in each of these components occur independently, their models can be

described as follows: Let the discrete random variables $G_{\rm B}$, $G_{\rm HT}$, $G_{\rm RP}$, and $G_{\rm LT}$, with universal sets specified by

$$\begin{aligned} &G_{\rm B} \in \Sigma_{G_{\rm B}} := \{{\rm N}_{\rm G_B}, {\rm A}_{\rm G_B}\},\ G_{\rm HT} \in \Sigma_{G_{\rm HT}} := \{{\rm N}_{\rm G_{HT}}, {\rm A}_{\rm G_{HT}}\}, \\ &G_{\rm RP} \in \Sigma_{G_{\rm RP}} := \{{\rm N}_{\rm G_{RP}}, {\rm A}_{\rm G_{RP}}\},\ G_{\rm LT} \in \Sigma_{G_{\rm LT}} := \{{\rm N}_{\rm G_{LT}}, {\rm A}_{\rm G_{LT}}\}, \end{aligned} \tag{25}$$

describe the 'health', i.e., normal or anomalous, of B, HT, RP, and LT, respectively. Under this scenario, we introduce the plant model for B as the conditional pmf $P(V_1|G_{\rm B})$. Regarding HT, as it follows from the independence of component anomalies, the plant model is given by $P(V_2 \in \{{\rm VL}_{\rm HT}, {\rm L}_{(2){\rm HT}}\}|G_{\rm HT})$ and $P(V_2 \in \{{\rm L}_{(1){\rm HT}}, {\rm N}_{\rm HT}\}|G_{\rm HT})$. To simplify notations, define the random variable V_2', whose universal set and pmf are

$$V_2' \in \{{\rm L}'_{\rm HT}, {\rm N}'_{\rm HT}\} \tag{26}$$

and

$$\begin{aligned} P(V_2' = {\rm L}'_{\rm HT}) &= P(\{V_2 = {\rm VL}_{\rm HT}\} \cup \{V_2 = {\rm L}_{(2){\rm HT}}\}), \\ P(V_2' = {\rm N}'_{\rm HT}) &= P(\{V_2 = {\rm L}_{(1){\rm HT}}\} \cup \{V_2 = {\rm N}_{\rm HT}\}), \end{aligned} \tag{27}$$

respectively. Utilizing (26) and (27), the plant model for HT can be re-expressed as $P(V_2'|G_{\rm HT})$. Similarly, for the other components, let

$$V_3' \in \{{\rm L}'_{\rm RP}, {\rm N}'_{\rm RP}\},\ V_4' \in \{{\rm N}'_{\rm LT}, {\rm H}'_{\rm LT}\}, \tag{28}$$

where

$$\begin{aligned} P(V_3' = {\rm L}'_{\rm RP}) &= P(\{V_3 = {\rm VL}_{\rm RP}\} \cup \{V_3 = {\rm L}_{(2){\rm RP}}\}), \\ P(V_3' = {\rm N}'_{\rm RP}) &= P(\{V_3 = {\rm L}_{(1){\rm RP}}\} \cup \{V_3 = {\rm N}_{\rm RP}\}) \end{aligned} \tag{29}$$

and

$$\begin{aligned} P(V_4' = {\rm N}'_{\rm LT}) &= P(\{V_4 = {\rm VL}_{(1){\rm LT}}\} \cup \{V_4 = {\rm L}_{(1){\rm LT}}\} \cup \{V_4 = {\rm M}_{(1){\rm LT}}\} \cup \{V_4 = {\rm N}_{\rm LT}\}), \\ P(V_4' = {\rm H}'_{\rm LT}) &= P(\{V_4 = {\rm VL}_{(2){\rm LT}}\} \cup \{V_4 = {\rm L}_{(2){\rm LT}}\} \cup \{V_4 = {\rm M}_{(2){\rm LT}}\} \cup \{V_4 = {\rm H}_{\rm LT}\}). \end{aligned} \tag{30}$$

Under (28)-(30) and the independence of component anomalies, the plant model for RP and LT are described by $P(V_3'|G_{\rm RP})$ and $P(V_4'|G_{\rm LT})$, respectively. Thus, the overall plant model is given by the following vector of conditional pmf's:

$$\left[P(V_1|G_{\rm B}),\ P(V_2'|G_{\rm HT}),\ P(V_3'|G_{\rm RP}),\ P(V_4'|G_{\rm LT})\right]. \tag{31}$$

The plant condition is expressed as the vector $\vec{G}$, specified by

$$\vec{G} \triangleq \left[P(G_{\rm B}),\ P(G_{\rm HT}),\ P(G_{\rm RP}),\ P(G_{\rm LT})\right], \tag{32}$$

where the pmf's in (32) are evaluated in Section VIII by utilizing (25)-(31) and assessments of the states of process variables.

F. Measure of resiliency

To quantify the accuracy of plant assessment, we introduce a *measure of resiliency* (*MR*) based on *Kullback-Leibler* (KL) *divergence* (see [10]). This divergence is typically used to quantify the 'distance' between two pmf's. For example, KL divergence between the actual health of B, denoted by $P(G_{\rm B})$, and the assessed health of B, denoted by $\hat{P}(G_{\rm B})$, is calculated as

$$D\left(P(G_{\rm B})||\hat{P}(G_{\rm B})\right) = \sum_{\sigma\in\Sigma_{G_{\rm B}}} P\{G_{\rm B}=\sigma\}\log_{|\Sigma_{G_{\rm B}}|}\frac{P\{G_{\rm B}=\sigma\}}{\hat{P}\{G_{\rm B}=\sigma\}}. \tag{33}$$

Utilizing the above divergence, we quantify MR for B as a normalized difference between $D\left(P(G_{\rm B})||\hat{P}_{\rm nr}(G_{\rm B})\right)$ and $D\left(P(G_{\rm B})||\hat{P}(G_{\rm B})\right)$, where $\hat{P}_{\rm nr}(G_{\rm B})$ is the estimated pmf of the *non-resilient system*, i.e., when the monitoring system uses the measurements of all sensors assuming $DQ_{ij}=1$, $\forall i,j$. In other words,

$$MR_{\rm B} = \frac{D\left(P(G_{\rm B})||\hat{P}_{\rm nr}(G_{\rm B})\right) - D\left(P(G_{\rm B})||\hat{P}(G_{\rm B})\right)}{D\left(P(G_{\rm B})||\hat{P}_{\rm nr}(G_{\rm B})\right)}. \tag{34}$$

Clearly, $MR_{\rm B}\leq 1$, and the value 1 is attained when $\hat{P}(G_{\rm B})=P(G_{\rm B})$. Similarly, MR can be calculated for HT, RP, and LT, which results in the following vector:

$$\overrightarrow{MR} = \left[MR_{\rm B},\ MR_{\rm HT},\ MR_{\rm RP},\ MR_{\rm LT}\right]. \tag{35}$$

Equations (33)-(35) are used in Section IX to quantify the efficacy of the monitoring system presented in this work.

Thus, the model of the power plant is characterized by models of process variables, serial connections, sensors, sensor subnetworks, plant, and the attacker.

IV. Data Quality Assessment Layer

As stated previously, sensor DQ is assigned using an active identification procedure developed in [2], which assumes knowledge of the d.c. gains of process variables at mid-points of their respective regions. For example, consider $\mathbf{V}_1$. The following d.c. gains are assumed to be known, and can be utilized for DQ evaluation:

$$\begin{aligned}\bar{\alpha}_{\rm L_B} &\triangleq \alpha_1\left(V^{\rm B}_{\rm min}+\tfrac{R^{\rm B}-V^{\rm B}_{\rm min}}{2}\right),\\ \bar{\alpha}_{\rm N_B} &\triangleq \alpha_1\left(R^{\rm B}+\tfrac{V^{\rm B}_{\rm max}-R^{\rm B}}{2}\right),\end{aligned} \tag{36}$$

where $\alpha_1(.)$ is defined in (5) and $V^{\rm B}_{\rm min}$, $R^{\rm B}$, and $V^{\rm B}_{\rm max}$ are defined in (2). According to the aforementioned procedure, probing signals (discussed in Section I) are simultaneously applied to each process variable using appropriate valves, and observed sensor responses are analyzed from the point of view of their

consistency with known d.c. gains (e.g., (36)). The sensors with larger consistency are viewed as more trustworthy, and their *DQ* is assigned accordingly. For brevity, the methodology to ascertain DQ is described below for sensors monitoring $\mathbf{V}_1$. Extension to sensors monitoring other process variables is straightforward.

In general, any type of deterministic or random probing signals could be used. Here, utilizing the fuel valve, we generate the simplest probe – a rectangular pulse of amplitude $A_0^{V_1}$ and duration T_1, applied at the time instant t_0, i.e.,

$$u_1(t) = A_0^{V_1}\text{rect}_{T_1}(t - t_0). \tag{37}$$

The value of $A_0^{V_1}$ is selected sufficiently small so that $\mathbf{V}_1$ remains in the same state (L_B or N_B) before and during the probe. Additionally, the small amplitude of the probe ensures that, while appropriately influencing $\mathbf{V}_1$ and $\mathbf{V}_3$, its effect on $\mathbf{V}_2$ and $\mathbf{V}_4$ is negligible. The value of T_1 is selected so that $\mathbf{V}_1$ reaches a small vicinity of its steady state defined by the probe.

Let the mean value, $E(\tilde{V}_1)$, of the process variable $\mathbf{V}_1$ be $\mu_{\tilde{V}_1}$. Further, let the mean value, $E(\tilde{S}_{1j})$, $j \in \{1, 2\}$, of the measurements of sensor $\mathbf{S}_{1j}$ before the probe be $\mu_{\tilde{S}_{1j}}$ and at the end of the probe be $\mu'_{\tilde{S}_{1j}}$. If the sensor is not compromised and its mean prior to the probe coincides exactly with that of the process variable, i.e., $\mu_{\tilde{S}_{1j}} = \mu_{\tilde{V}_1}$, the following holds true:

$$\begin{aligned} \mu'_{\tilde{S}_{1j}} - \mu_{\tilde{S}_{1j}} &= A_0^{V_1}\alpha_1(\mu_{\tilde{V}_1}), \\ &= A_0^{V_1}\alpha_1(\mu_{\tilde{S}_{1j}}), \end{aligned} \tag{38}$$

where $\alpha_1(.)$, as mentioned before, is defined in (5). However, if the sensor is compromised, the previous statement does not hold true. In order to distinguish between a captured sensor and one that is not, we introduce *probing inconsistency* (PIC) of $\mathbf{S}_{1j}$ as follows:

$$PIC_{1j} \stackrel{\triangle}{=} \left|(\mu'_{\tilde{S}_{1j}} - \mu_{\tilde{S}_{1j}}) - A_0^{V_1}\bar{\alpha}_1\left(\mu_{\tilde{S}_{1j}}\right)\right|, \tag{39}$$

where $\bar{\alpha}_1(\mu_{\tilde{S}_{1j}})$ is defined as

$$\bar{\alpha}_1(\mu_{\tilde{S}_{1j}}) = \begin{cases} \bar{\alpha}_{\text{L}_\text{B}}, & \text{if } \mu_{\tilde{S}_{1j}} \in [V_{\min}^{\text{B}}, R^{\text{B}}) \\ \bar{\alpha}_{\text{N}_\text{B}}, & \text{if } \mu_{\tilde{S}_{1j}} \in [R^{\text{B}}, V_{\max}^{\text{B}}]. \end{cases} \tag{40}$$

The d.c. gains $\bar{\alpha}_{\text{L}_\text{B}}$ and $\bar{\alpha}_{\text{N}_\text{B}}$, involved in (40), are defined in (36). Clearly, PIC_{1j} is large if $\mathbf{S}_{1j}$ is captured and is relatively small if not.

Given the above PIC, data quality is assigned according to

$$DQ_{1j} = e^{-F(PIC_{1j})}, \tag{41}$$

where $F(.)$ is a monotonically increasing function of PIC, with $F(0) = 0$. It turns out that a convenient way of introducing this function is

$$F(PIC_{1j}) \;\; = \;\; -\frac{\ln \epsilon}{PIC_{\mathrm{M}_1}^2} PIC_{1j}^2, \tag{42}$$

where $\epsilon << 1$ (a design parameter) is the DQ assigned when PIC is PIC_{M_1}, given by

$$PIC_{\mathrm{M}_1} = |A_0^{V_1} \cdot (\bar{\alpha}_{\mathrm{L_B}} - \bar{\alpha}_{\mathrm{N_B}})|. \tag{43}$$

Equations (37)-(43) characterize the DQ assessment layer of the power plant monitoring system.

V. Process Variable Assessment Layer

As illustrated in Figure 4, the process variable assessment layer evaluates the joint pmf's of process variables in each subplant. These calculations are based on the previously mentioned h-procedure, Dempster-Shafer combination rule, and intra-subplant inferences. The h-procedure is a modified stochastic approximation algorithm developed in [3], and is utilized to calculate the pmf of a process variable based on the measurements of one sensor that monitors the variable. Using this procedure, multiple estimates of the process variable pmf can be calculated by utilizing the measurements of several sensors. The Dempster-Shafer rule is used to combine these pmf's into a single pmf. Given two serially connected process variables of a subplant, the pmf of one can be calculated based on the measurements of a sensor that monitors the other one. This computation, termed as intra-subplant inference, is based on statistical models of the serial connection and, once again, the h-procedure and the Dempster-Shafer rule. Details of these calculations are provided below.

A. Process variable pmf estimation for subplant $\mathbf{G}_{\mathrm{I}}$

1) PMF estimation based on a single sensor monitoring the process variable: As mentioned before, process variable $\mathbf{V}_1$ is monitored by two sensors, namely, $\mathbf{S}_{11}$ and $\mathbf{S}_{12}$. Consider sensor $\mathbf{S}_{11}$, and let its data quality be DQ_{11}. Further, let $s_1^{(11)}, s_2^{(11)}, ..., s_n^{(11)}$ denote the sensor's measurements up to time instant n and let $\hat{P}_n^{S_{11}}(V_1)$ be the estimate of the pmf of V_1 based on these measurements and DQ_{11}. In other words,

$$\hat{P}_n^{S_{11}}(V_1 = \sigma_1) \stackrel{\triangle}{=} P(V_1 = \sigma_1 | s_1^{(11)}, s_2^{(11)}, ..., s_n^{(11)}; DQ_{11}), \tag{44}$$

where $n \in \mathbb{N}$ and $\sigma_1 \in \Sigma_1$. Our goal is to determine the pmf $\hat{P}^{S_{11}}(V_1)$, defined as

$$\hat{P}^{S_{11}}(V_1 = \sigma_1) \stackrel{\triangle}{=} \lim_{n \to \infty} P(V_1 = \sigma_1 | s_1^{(11)}, s_2^{(11)}, ..., s_n^{(11)}; DQ_{11}). \tag{45}$$

For convenience, let us denote $\hat{P}_n^{S_{11}}(V_1 = \sigma_1)$ as $h_{\sigma_1}(n)$. With ϵ_h being a design parameter and the initial conditions specified as $h_{\sigma_1}(0) = \frac{1}{|\Sigma_1|} = \frac{1}{2}$, $\forall \sigma_1 \in \Sigma_1$, we introduce the following recursive procedure for calculating $h_{\sigma_1}(n)$:

$$h_{\sigma_1}(n+1) = h_{\sigma_1}(n) + \epsilon_h \left[h^*_{\sigma_1}\left(s_{n+1}^{(11)}\right) - h_{\sigma_1}(n) \right]. \tag{46}$$

Regarding the set point of (46), i.e., $h^*_{\sigma_1}\left(s_{n+1}^{(11)}\right)$, it is defined, based on the sensor believability (20), as follows:

$$h^*_{\sigma_1}\left(s_{n+1}^{(11)}\right) = \begin{cases} \beta_{11}, & \text{if } s_{n+1}^{(11)} = \sigma_1 \\ \frac{1-\beta_{11}}{|\Sigma_1|-1}, & \text{if } s_{n+1}^{(11)} \neq \sigma_1. \end{cases} \tag{47}$$

Equations (46) and (47) characterize the h-procedure.

It was shown in [3] that as $n \to \infty$, the h-procedure converges in probability to the following limit:

$$h_{\sigma_1}(n) \xrightarrow{\text{P}} DQ_{11} \cdot P(S_{11} = \sigma_1) + \frac{1 - DQ_{11}}{2},\ \sigma_1 \in \Sigma_1. \tag{48}$$

Clearly, according to (48), the estimate $\hat{P}^{S_{11}}(V_1)$ depends not only on the sensor measurements, but also on DQ_{11}. Observe that if DQ_{11} is close to 1, the estimated pmf of V_1 is close to the pmf of S_{11}, which is identical to what is postulated by classical statistics. However, if DQ_{11} is close to 0, the same measurements result in $\hat{P}^{S_{11}}(V_1)$ being practically uniform and independent of the sensor measurements. For all intermediate values of DQ_{11}, the estimate $\hat{P}^{S_{11}}(V_1)$ is an affine function of DQ_{11}.

The corresponding estimate of the pmf of V_1 based on the measurements of sensor $\mathbf{S}_{12}$ is calculated in the same manner as above.

As before, let $\hat{P}^{S_{21}}(V_2)$ and $\hat{P}^{S_{22}}(V_2)$ denote the estimates of the pmf of V_2 based on measurements of sensors $\mathbf{S}_{21}$ and $\mathbf{S}_{22}$, respectively. These pmf's are also calculated using the h-procedure described above.

2) PMF estimation based on multiple sensors monitoring a single process variable: Let $\hat{P}^{S_{11}S_{12}}(V_1)$ denote the estimate of the pmf of V_1 based on the measurements of both sensors that monitor the process variable, namely, $\mathbf{S}_{11}$ and $\mathbf{S}_{12}$. In other words,

$$\hat{P}^{S_{11}S_{12}}(V_1) \stackrel{\triangle}{=} \lim_{n\to\infty} P(V_1 | s_1^{(11)}, ..., s_n^{(11)}; DQ_{11}; s_1^{(12)}, ..., s_n^{(12)}; DQ_{12}). \tag{49}$$

To obtain the sought estimate, we combine the two pmf's $\hat{P}^{S_{11}}(V_1)$ and $\hat{P}^{S_{12}}(V_1)$, calculated above, into a single pmf using Dempster-Shafer rule (see [9]):

$$\hat{P}^{S_{11}S_{12}}(V_1 = \sigma_1) = \frac{\hat{P}^{S_{11}}(V_1 = \sigma_1)\hat{P}^{S_{12}}(V_1 = \sigma_1)}{\sum_{\sigma_1} \hat{P}^{S_{11}}(V_1 = \sigma_1)\hat{P}^{S_{12}}(V_1 = \sigma_1)},\ \sigma_1 \in \Sigma_1. \tag{50}$$

Regarding process variable $\mathbf{V}_2$, the pmf $\hat{P}^{S_{21}S_{22}}(V_2)$ is calculated in the same manner as above.

3) Estimation of process variable pmf's based on intra-subplant inferences: The two goals of these calculations are as follows: One, as mentioned before, is to compute estimates of process variable pmf's based on the serial connection of the subplant. The other one is to compute estimates of the joint pmf of process variables, which are utilized in the adaptation/rational controller layer as the basis for adaptation. Regarding subplant $\mathbf{G}_{\mathrm{I}}$, the serial connection is characterized by the conditional pmf's $P(V_1|V_2)$ and $P(V_2|V_1)$, introduced in (16) and (17), respectively.

Let $\hat{P}^{S_{21}}(V_1)$ denote the estimate of the pmf of V_1 based on the measurements of a sensor that monitors $\mathbf{V}_2$, namely, $\mathbf{S}_{21}$. In other words,

$$\hat{P}^{S_{21}}(V_1) \stackrel{\triangle}{=} \lim_{n\to\infty} P(V_1|s_1^{(21)},...,s_n^{(21)};DQ_{21}). \tag{51}$$

To calculate the sought pmf, we first evaluate $\hat{P}^{S_{21}}(V_2)$ using the h-procedure. Next, we utilize $P(V_1|V_2)$ and the total probability formula to compute $\hat{P}^{S_{21}}(V_1)$, i.e.,

$$\hat{P}^{S_{21}}(V_1) = \sum_{\sigma_2} P(V_1|V_2=\sigma_2)\hat{P}^{S_{21}}(V_2=\sigma_2),\ \sigma_2 \in \Sigma_2. \tag{52}$$

Similarly, the pmf $\hat{P}^{S_{22}}(V_1)$ can be calculated using the procedure described above.

Let $\hat{P}^{S_{11}S_{21}}(V_1)$ denote the estimate of the pmf of V_1 based on measurements of two sensors, namely, $\mathbf{S}_{11}$ and $\mathbf{S}_{21}$, which monitor $\mathbf{V}_1$ and $\mathbf{V}_2$, respectively. In this scenario, $\hat{P}^{S_{11}S_{21}}(V_1)$ is obtained by combining $\hat{P}^{S_{11}}(V_1)$ and $\hat{P}^{S_{21}}(V_1)$ using, once again, the Dempster-Shafer rule. This procedure can be extended to compute the pmf of V_1 based on any other combination of sensors. The same calculations are applied to evaluate the pmf's of V_2 as well.

Regarding the joint pmf of V_1 and V_2, it is calculated by utilizing the inference model $P(V_1|V_2)$ and the estimate of the pmf of V_2. For instance, the joint pmf $\hat{P}^{S_{11}}(V_1,V_2)$ is computed as

$$\hat{P}^{S_{11}}(V_1,V_2) = P(V_1|V_2)\hat{P}^{S_{11}}(V_2). \tag{53}$$

Regarding subplants $\mathbf{G}_{\mathrm{II}}$ and $\mathbf{G}_{\mathrm{III}}$, the process variable pmf identification procedure is identical to the one described above. The corresponding joint pmf's of process variables are calculated as in (53). For instance, the pmf's $\hat{P}^{S_{32}}(V_1,V_3)$ and $\hat{P}^{S_{41}S_{42}}(V_3,V_4)$ are computed as

$$\begin{aligned} \hat{P}^{S_{32}}(V_1,V_3) &= P(V_1|V_3)\hat{P}^{S_{32}}(V_3), \\ \hat{P}^{S_{41}S_{42}}(V_3,V_4) &= P(V_3|V_4)\hat{P}^{S_{41}S_{42}}(V_4), \end{aligned} \tag{54}$$

where the inference models $P(V_1|V_3)$ and $P(V_3|V_4)$ are given in Subsection III-B.

B. Example of intra-subplant inference

In this example, we illustrate the utility of intra-subplant inference by quantifying the loss of information when it is not applied.

Consider subplant $\mathbf{G}_{\mathrm{III}}$. Assume that RP is undamaged, while LT is malfunctioning, i.e., $\mathbf{V}_3$ and $\mathbf{V}_4$ are in states N_{RP} and H_{LT}, respectively. Further, assume that only sensor $\mathbf{S}_{41}$, which monitors $\mathbf{V}_4$, is utilized to ascertain the states of both RP and LT. Let the pmf $\hat{P}^{S_{41}}(V_4)$ be specified as $[0,0,0,0,0,0,0,1]$, where the row vector indicates probabilities of $\mathbf{V}_4$ being in states $\mathrm{VL}_{(1)\mathrm{LT}}$, $\mathrm{VL}_{(2)\mathrm{LT}}$, $\mathrm{L}_{(1)\mathrm{LT}}$, $\mathrm{L}_{(2)\mathrm{LT}}$, $\mathrm{M}_{(1)\mathrm{LT}}$, $\mathrm{M}_{(2)\mathrm{LT}}$, N_{LT}, and H_{LT}. Given these data, the problem is to assess the pmf of V_3.

Scenario 1: Utilizing intra-subplant inference,

$$\hat{P}^{S_{41}}(V_3) = \sum_{\sigma_4 \in \Sigma_4} P(V_3|V_4 = \sigma_4)\hat{P}^{S_{41}}(V_4 = \sigma_4), \tag{55}$$

we obtain $\hat{P}^{S_{41}}(V_3) = [0,0,0,1]$, which indicates accurately that $\mathbf{V}_3$ is in state N_{RP}.

Scenario 2: Here, we ascertain the pmf of V_3, denoted as $\hat{P}(V_3)$, without utilizing inferences. Clearly, we possess no knowledge, whatsoever, about the state of $\mathbf{V}_3$, since only measurements of $\mathbf{S}_{41}$ are taken into account. Therefore, $\hat{P}(V_3)$ is assigned as $\left[\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{4}\right]$, i.e., all states of $\mathbf{V}_3$ are equiprobable.

Clearly, these two scenarios illustrate the efficacy of intra-subplant inferences for process variable assessment.

VI. Adaptation/Rational Controller Layer

As mentioned in Section II, the adaptation/rational controller (A/RC) layer is based on rational controllers developed in [4]–[6]. In this section, we describe the operation of these controllers and specify their parameters and temporal properties.

A. Rational controllers and subplant process variables pmf assessment

Each subnetwork is equipped with a decision making device, namely, a rational controller, which has two properties: *ergodicity* and *rationality*. The ergodicity property implies that all states in the decision space are visited with a non-zero probability. The rationality property implies that the residence time in states with a smaller penalty function is larger than in those with a larger one. The degree to which this distinction takes place is referred to as the *level of rationality*.

Regarding subplant $\mathbf{G}_{\mathrm{I}}$, the penalty function of the rational controller is chosen as

$$\Phi_{\mathrm{I}}(x_{\mathrm{I}}) = I\left\{\hat{P}_{x_{\mathrm{I}}}(V_1, V_2)\right\}, \tag{56}$$

where $\hat{P}_{x_{\rm I}}(V_1, V_2)$ is the estimate of the joint pmf of V_1 and V_2 associated with sensor subnetwork state $x_{\rm I} \in X_{\rm I}$ (see (23)) and $I\{.\}$ is the entropy of the pmf. For example, the entropy of $P(V_1)$ is calculated as

$$I\left\{P(V_1)\right\} = -\sum_{\sigma_1 \in \Sigma_1} P(V_1 = \sigma_1) \log_{|\Sigma_1|} P(V_1 = \sigma_1). \tag{57}$$

The residence time, $T_{x_{\rm I}}$, for the rational controller is specified as follows:

$$T_{x_{\rm I}} = \begin{cases} T_{\max}, & \text{if } \Phi_{\rm I}(x_{\rm I}) \leq \eta \\ \left(\frac{\eta}{\Phi_{\rm I}(x_{\rm I})}\right)^N T_{\max}, & \text{if } \Phi_{\rm I}(x_{\rm I}) > \eta, \end{cases} \tag{58}$$

where N (level of rationality $-$ a positive integer), $\eta > 0$ (a small number), and $T_{\max}$ (the largest residence time) are design parameters. It is clear from (56) and (58) that the rational controller resides longer in states of $X_{\rm I}$ where the process variable pmf's possess smaller entropy, i.e., less uncertainty. Further, to ensure ergodicity, the rational controller visits all states of the sensor subnetwork in a deterministic, round-robin manner.

Denote the relative residence time of the rational controller in state $x_{\rm I} \in X_{\rm I}$ after the controller visits all states of $X_{\rm I}$ as $\tau_{x_{\rm I}}$:

$$\tau_{x_{\rm I}} = \frac{T_{x_{\rm I}}}{\sum_{x_{\rm I} \in X_{\rm I}} T_{x_{\rm I}}}. \tag{59}$$

Then, based on $\tau_{x_{\rm I}}$, the process variable pmf's associated with subplant $\mathbf{G}_{\rm I}$ can be calculated as follows:

$$\begin{aligned} \hat{P}_{\mathbf{G}_{\rm I}}(V_1) &= \sum_{x_{\rm I} \in X_{\rm I}} \tau_{x_{\rm I}} \hat{P}_{x_{\rm I}}(V_1), \\ \hat{P}_{\mathbf{G}_{\rm I}}(V_2) &= \sum_{x_{\rm I} \in X_{\rm I}} \tau_{x_{\rm I}} \hat{P}_{x_{\rm I}}(V_2). \end{aligned} \tag{60}$$

Similarly, the pmf's $\hat{P}_{\mathbf{G}_{\rm II}}(V_1)$ and $\hat{P}_{\mathbf{G}_{\rm II}}(V_3)$, associated with $\mathbf{G}_{\rm II}$, and $\hat{P}_{\mathbf{G}_{\rm III}}(V_3)$ and $\hat{P}_{\mathbf{G}_{\rm III}}(V_4)$, associated with $\mathbf{G}_{\rm III}$, can be evaluated by utilizing rational controllers as in (56) and (58).

The pmf's $\hat{P}_{\mathbf{G}_{\rm I}}(V_1)$, $\hat{P}_{\mathbf{G}_{\rm II}}(V_1)$, $\hat{P}_{\mathbf{G}_{\rm I}}(V_2)$, $\hat{P}_{\mathbf{G}_{\rm II}}(V_3)$, $\hat{P}_{\mathbf{G}_{\rm III}}(V_3)$, and $\hat{P}_{\mathbf{G}_{\rm III}}(V_4)$ are used in Section VII to combine information from the three subnetworks and obtain pmf's of process variables that are subsequently utilized to evaluate the condition of the power plant.

B. Temporal properties of adaptation

From the temporal point of view, the three rational controllers described in Subsection VI-A are assumed to adapt simultaneously over their respective subnetworks. In this scenario, the A/RC layer is said to

consist of *epochs*; For each subnetwork state $x_k \in X_k$, $k \in \{\mathrm{I}, \mathrm{II}, \mathrm{III}\}$, the epoch consists of three periods:

DQ evaluation period, T_{DQ},

Process variable pmf evaluation period, T_{eval},

Residence period in state x_k, T_{x_k}.

If Q is the number of states in the subnetwork, then Q epochs comprise a *cycle*. At the end of each cycle, the process variable pmf's are identified and can be utilized for inter-subplant inferences.

Assuming that the sensor measurements are provided every 0.01sec, the procedure described in Section IV is considered to be of duration $T_{DQ} = 5$sec. Using the procedures described in Section V, duration of process variable assessment, T_{eval}, is about 6sec. The maximum residence period, T_{max}, can be selected as desired. If T_{max} is selected to be 1sec, the duration of each epoch is less than or equal to 12sec.

As mentioned above, $Q = 15$ epochs constitute a cycle, wherein each of Q states is visited. So, for each subnetwork, the cycle duration is, at most, $12Q = 180$sec. Moreover, every rational controller holds its pmf's until the others have completed their respective cycles. Thus, the A/RC layer evaluates the process variable pmf's of the three subplants within, at most, 180sec.

VII. Knowledge fusion layer

In this layer, the pmf's $\hat{P}_{\mathbf{G}_{\mathrm{I}}}(V_1)$, $\hat{P}_{\mathbf{G}_{\mathrm{II}}}(V_1)$, $\hat{P}_{\mathbf{G}_{\mathrm{I}}}(V_2)$, $\hat{P}_{\mathbf{G}_{\mathrm{II}}}(V_3)$, $\hat{P}_{\mathbf{G}_{\mathrm{III}}}(V_3)$, and $\hat{P}_{\mathbf{G}_{\mathrm{III}}}(V_4)$, evaluated at the A/RC layer, are combined using inter-subplant inferences to obtain $\hat{P}^*(V_1)$, $\hat{P}^*(V_2)$, $\hat{P}^*(V_3)$, and $\hat{P}^*(V_4)$. This is described below.

A. Calculations

We infer the state of $\mathbf{V}_1$ based on the state of $\mathbf{V}_3$ from subplant $\mathbf{G}_{\mathrm{III}}$ using total probability formula:

$$\hat{P}_{\mathbf{G}_{\mathrm{III}}}(V_1) = \sum_{\sigma_3 \in \Sigma_3} P(V_1 | V_3 = \sigma_3) \hat{P}_{\mathbf{G}_{\mathrm{III}}}(V_3 = \sigma_3), \tag{61}$$

where Σ_3 is defined in (11). Next, we combine $\hat{P}_{\mathbf{G}_{\mathrm{I}}}(V_1)$, $\hat{P}_{\mathbf{G}_{\mathrm{II}}}(V_1)$, and $\hat{P}_{\mathbf{G}_{\mathrm{III}}}(V_1)$ using Dempster-Shafer rule to obtain $\hat{P}_{\mathbf{G}_{\mathrm{I,II,III}}}(V_1)$:

$$\hat{P}_{\mathbf{G}_{\mathrm{I,II,III}}}(V_1 = \sigma_1) = \frac{\prod_{k=\mathrm{I}}^{\mathrm{III}} \hat{P}_{\mathbf{G}_k}(V_1 = \sigma_1)}{\sum_{\sigma_1 \in \Sigma_1} \prod_{k=\mathrm{I}}^{\mathrm{III}} \hat{P}_{\mathbf{G}_k}(V_1 = \sigma_1)}, \tag{62}$$

where Σ_1 is defined in (3).

Based on $\left\{\hat{P}_{\mathbf{G}_{\rm I}}(V_1), \hat{P}_{\mathbf{G}_{\rm II}}(V_1), \hat{P}_{\mathbf{G}_{\rm III}}(V_1), \hat{P}_{\mathbf{G}_{\rm I,II,III}}(V_1)\right\}$, we select $\hat{P}^*(V_1)$ as one of them that possesses the smallest entropy. In other words, if $I\{P(.)\}$ denotes the entropy of pmf $P(.)$, then

$$\hat{P}^*(V_1) = \arg\min\left\{I\left\{\hat{P}_{\mathbf{G}_{\rm I}}(V_1)\right\}, I\left\{\hat{P}_{\mathbf{G}_{\rm II}}(V_1)\right\}, I\left\{\hat{P}_{\mathbf{G}_{\rm III}}(V_1)\right\}, I\left\{\hat{P}_{\mathbf{G}_{\rm I,II,III}}(V_1)\right\}\right\}. \tag{63}$$

As for $\mathbf{V}_2$, we infer its pmf based, again, on total probability formula:

$$\hat{P}_{V_1^*}(V_2) = \sum_{\sigma_1 \in \Sigma_1} P(V_2|V_1 = \sigma_1)\hat{P}^*(V_1 = \sigma_1). \tag{64}$$

Next, we combine the pmf's $\hat{P}_{V_1^*}(V_2)$ and $\hat{P}_{\mathbf{G}_{\rm I}}(V_2)$ using Dempster-Shafer rule to obtain $\hat{P}_{\mathbf{G}_{\rm I},V_1^*}(V_2)$:

$$\hat{P}_{\mathbf{G}_{\rm I},V_1^*}(V_2 = \sigma_2) = \frac{\hat{P}_{\mathbf{G}_{\rm I}}(V_2 = \sigma_2)\hat{P}_{V_1^*}(V_2 = \sigma_2)}{\sum\limits_{\sigma_2 \in \Sigma_2} \hat{P}_{\mathbf{G}_{\rm I}}(V_2 = \sigma_2)\hat{P}_{V_1^*}(V_2 = \sigma_2)}, \tag{65}$$

where Σ_2 is defined in (7). Finally, based on $\left\{\hat{P}_{\mathbf{G}_{\rm I}}(V_2), \hat{P}_{V_1^*}(V_2), \hat{P}_{\mathbf{G}_{\rm I},V_1^*}(V_2)\right\}$, we select $\hat{P}^*(V_2)$ as one of them that possesses the smallest entropy, i.e.,

$$\hat{P}^*(V_2) = \arg\min\left\{I\left\{\hat{P}_{\mathbf{G}_{\rm I}}(V_2)\right\}, I\left\{\hat{P}_{V_1^*}(V_2)\right\}, I\left\{\hat{P}_{\mathbf{G}_{\rm I},V_1^*}(V_2)\right\}\right\}. \tag{66}$$

The inferences about $\mathbf{V}_3$ and $\mathbf{V}_4$ are obtained in a manner similar to that of $\mathbf{V}_1$ and $\mathbf{V}_2$, respectively. These calculations yield the pmf's $\hat{P}^*(V_3)$ and $\hat{P}^*(V_4)$.

Thus, the knowledge fusion layer provides estimates of the pmf's of all process variables to be used in the subsequent plant assessment layer to evaluate the condition of B, HT, RP, and LT.

B. Accuracy

Obviously, the knowledge fusion layer combines the information obtained within the overlapping decomposition $\{\mathbf{G}_{\rm I}, \mathbf{G}_{\rm II}, \mathbf{G}_{\rm III}\}$ in order to obtain the pmf's of all process variables, $\mathbf{V}_1$, $\mathbf{V}_2$, $\mathbf{V}_3$, and $\mathbf{V}_4$. These pmf's can, in principle, be obtained using the centralized system of Figure **??**. The following question arises: How well do the pmf's $\hat{P}^*(V_i)$, $i = 1, 2, 3, 4$, approximate the ones that can be obtained using the centralized system and denoted as $\hat{P}^*_{\rm c}(V_i)$, $i = 1, 2, 3, 4$. Although at present we do not have a complete answer to this question, below we analyze a simplified problem that provides a sufficient condition when the pmf's $\hat{P}^*(V_i)$ and $\hat{P}^*_{\rm c}(V_i)$ are the same.

Let $\mathbf{V}_1$ and $\mathbf{V}_2$ be two serially connected process variables (similar, but not identical, to the ones introduced in Section III). Assume that the universal sets of V_1 and V_2 are

$$\begin{aligned} \Sigma_1 &\stackrel{\triangle}{=} \{\mathrm{N}_1, \mathrm{A}_1\}, \\ \Sigma_2 &\stackrel{\triangle}{=} \{\mathrm{N}_2, \mathrm{A}_{21}, \mathrm{A}_{22}, \mathrm{A}_{23}\}, \end{aligned} \tag{67}$$

where, as before, 'N' is Normal and 'A' is Anomaly. Let Z_{N_1} and Z_{A_1} be the intervals where $\mathbf{V}_1$ is viewed as being in states N_1 and A_1, respectively (see, for instance, (1)). Similarly, let Z_{N_2}, $Z_{\mathrm{A}_{21}}$, $Z_{\mathrm{A}_{22}}$, and $Z_{\mathrm{A}_{23}}$ be the corresponding intervals for $\mathbf{V}_2$. Assume that the d.c. gains of the process variables are piecewise constant in these intervals, i.e.,

$$\alpha_1(\tilde{V_1}) \;=\; \begin{cases} \alpha_{\mathrm{N}_1}, & \text{if } \tilde{V_1} \in Z_{\mathrm{N}_1} \\ \alpha_{\mathrm{A}_1}, & \text{if } \tilde{V_1} \in Z_{\mathrm{A}_1}, \end{cases} \tag{68}$$

and

$$\alpha_2(\tilde{V_2}) \;=\; \begin{cases} \alpha_{\mathrm{N}_2}, & \text{if } \tilde{V_2} \in Z_{\mathrm{N}_2} \\ \alpha_{\mathrm{A}_{21}}, & \text{if } \tilde{V_2} \in Z_{\mathrm{A}_{21}} \\ \alpha_{\mathrm{A}_{22}}, & \text{if } \tilde{V_2} \in Z_{\mathrm{A}_{22}} \\ \alpha_{\mathrm{A}_{23}}, & \text{if } \tilde{V_2} \in Z_{\mathrm{A}_{23}}. \end{cases} \tag{69}$$

Let the serial connection of $\mathbf{V}_1$ and $\mathbf{V}_2$ be characterized by the conditional pmf's $P(V_1|V_2)$ and $P(V_2|V_1)$, which are given by (16) and (17), respectively. Finally, assume that each process variable is monitored by two sensors: $\mathbf{S}_{11}$ and $\mathbf{S}_{12}$ monitor $\mathbf{V}_1$, while $\mathbf{S}_{21}$ and $\mathbf{S}_{22}$ monitor $\mathbf{V}_2$.

As before, introduce two subplants, $\mathbf{G}_{\mathrm{I}}$ and $\mathbf{G}_{\mathrm{II}}$, in order to combat the curse of dimensionality; $\mathbf{G}_{\mathrm{I}}$ consists of $\mathbf{V}_1$ and associated sensors, while $\mathbf{G}_{\mathrm{II}}$ comprises $\mathbf{V}_2$ and associated sensors. Two methods for evaluating pmf's of V_1 and V_2 are considered below, the centralized and the decentralized, and their results are compared.

Centralized process variable assessments (CPVA): Each sensor is assigned DQ based on the procedure described in Section IV. The state space of the overall sensor network is given by

$$X = \{(1000), (0100), ..., (1111)\}\,, \tag{70}$$

which contains $2^4 - 1 = 15$ states. Let the sensor network be equipped with a rational controller, whose objective is to minimize the following penalty function:

$$\Phi(x) = I\left\{\hat{P}_x(V_1, V_2)\right\},\ x \in X, \tag{71}$$

where $\hat{P}_x(V_1, V_2)$ is computed using the h-procedure (46), total probability formula (52), and Dempster -Shafer rule (50) (see Subsection V-A), and $I\{.\}$ denotes the entropy. Assume that a unique solution of this minimization problem exists and is given by x^*, i.e.,

$$x^* = \arg\min_{x \in X} \Phi(x). \tag{72}$$

To solve the above problem, we specify the residence time of the controller in state $x \in X$ as

$$T_x = \left(\frac{1}{\Phi(x)} \right)^N, \tag{73}$$

where N is sufficiently large. In this scenario, it has been shown in [4] that the rational controller determines the optimal state x^*. After x^* is ascertained, the pmf $\hat{P}_{x^*}(V_1, V_2)$ is marginalized to obtain $\hat{P}_{x^*}(V_1)$ and $\hat{P}_{x^*}(V_2)$.

Given that the above rational controller operates on X, we henceforth refer to it as the *global rational controller*.

Decentralized process variable assessments with inter-subplant inferences (DPVA-I): Assign DQ to each sensor as before. Decompose the sensor network with state space (70) into two subnetworks, with state spaces $X_{\rm I}$ and $X_{\rm II}$ defined as follows:

$$\begin{aligned} X_{\rm I} &= \{(10)_{\rm I}, (01)_{\rm I}, (11)_{\rm I}\}, \\ X_{\rm II} &= \{(10)_{\rm II}, (01)_{\rm II}, (11)_{\rm II}\}. \end{aligned} \tag{74}$$

Assume that each subnetwork is equipped with a *local rational controller*, whose objective is to minimize the penalty functions

$$\begin{aligned} \Phi_{\rm I}(x_{\rm I}) &= I\left\{\hat{P}_{x_{\rm I}}(V_1)\right\},\ x_{\rm I} \in X_{\rm I}, \\ \Phi_{\rm II}(x_{\rm II}) &= I\left\{\hat{P}_{x_{\rm II}}(V_2)\right\},\ x_{\rm II} \in X_{\rm II}, \end{aligned} \tag{75}$$

respectively, where $\hat{P}_{x_{\rm I}}(V_1)$ and $\hat{P}_{x_{\rm II}}(V_2)$ are calculated using the h-procedure (46) and Dempster-Shafer rule (50). Assume that unique solutions of these minimization problems exist and are given by $x^*_{\rm I}$ and $x^*_{\rm II}$:

$$\begin{aligned} x^*_{\rm I} &= \arg\min_{x_{\rm I} \in X_{\rm I}} \Phi_{\rm I}(x_{\rm I}), \\ x^*_{\rm II} &= \arg\min_{x_{\rm II} \in X_{\rm II}} \Phi_{\rm II}(x_{\rm II}). \end{aligned} \tag{76}$$

Let the residence time of the rational controllers be specified as in (73). Under this scenario, the local optima, $x^*_{\rm I}$ and $x^*_{\rm II}$, are determined, and the pmf's $\hat{P}_{x^*_{\rm I}}(V_1)$ and $\hat{P}_{x^*_{\rm II}}(V_2)$ are utilized for inter-subplant inferences as follows:

For $\mathbf{V}_2$, the pmf $\hat{P}_{x^*_{\rm I}}(V_2)$ is computed using total probability formula:

$$\hat{P}_{x^*_{\rm I}}(V_2) = \sum_{\sigma_1 \in \Sigma_1} P(V_2 | V_1 = \sigma_1) \hat{P}_{x^*_{\rm I}}(V_1 = \sigma_1). \tag{77}$$

Next, the pmf $\hat{P}_{(x^*_{\rm I}, x^*_{\rm II})}(V_2)$ is calculated by combining $\hat{P}_{x^*_{\rm I}}(V_2)$ and $\hat{P}_{x^*_{\rm II}}(V_2)$ using Dempster-Shafer rule. The pmf with smallest entropy among $\hat{P}_{x^*_{\rm I}}(V_2)$, $\hat{P}_{x^*_{\rm II}}(V_2)$, and $\hat{P}_{(x^*_{\rm I}, x^*_{\rm II})}(V_2)$ is reported to the operator.

Similar calculations are applied for $\mathbf{V}_1$ as well.

Remark 7.1: Note that X can be expressed in terms of the concatenation of subnetworks $X_{\rm I}$ and $X_{\rm II}$:

$$X = \{X_{\rm I} \times X_{\rm II}\} \cup \{X_{\rm I} \times (00)\} \cup \{(00) \times X_{\rm II}\}, \tag{78}$$

where $\times$ denotes the Cartesian product. □

Remark 7.2: Note that, as follows from the example of Subsection V-B, the performance of the decentralized system without inferences is inferior to that of the centralized one. □

Let V be an arbitrary discrete random variable with the universal set Σ. Assume that there exist two candidate pmf's, $P_1(V)$ and $P_2(V)$, of V. These two pmf's can be combined using Dempster-Shafer rule to obtain another pmf, $P_{12}(V)$, of V. Using these pmf's, we introduce:

Definition 7.1: The pmf's $P_1(V)$ and $P_2(V)$ are termed *Dempster-Shafer monotonic* (DS-monotonic) if the following two conditions hold:

$$\arg\max_{\sigma\in\Sigma} P_1(V=\sigma) = \arg\max_{\sigma\in\Sigma} P_2(V=\sigma), \tag{79}$$

$$I\{P_{12}(V)\} < \min\left[I\{P_1(V)\}, I\{P_2(V)\}\right]. \tag{80}$$

In other words, $P_1(V)$ and $P_2(V)$ are DS-monotonic if they indicate the same state with highest probability, and the pmf $P_{12}(V)$ has a smaller entropy than the constituent ones. This definition has an obvious extension for more than two pmf's as well.

Theorem 7.1: Assume that the pmf's $\hat{P}_{x_{\rm I}^*}(V_2)$ and $\hat{P}_{x_{\rm II}^*}(V_2)$ are DS-monotonic. Then, $(x_{\rm I}^*, x_{\rm II}^*) = x^*$, and the pmf's of V_1 and V_2 calculated using CPVA and DPVA-I are, respectively, identical.

Proof: See the Appendix. □

Based on the above theorem, we hypothesize: *The pmf's of process variables* $\mathbf{V}_1$, $\mathbf{V}_2$, $\mathbf{V}_3$, *and* $\mathbf{V}_4$ *identified using the centralized system of Figure* **??** *and the overlapping decomposition architecture of Figure 4 are identical if the following four groups of pmf's are DS-monotonic:* $\left\{\hat{P}_{\mathbf{G}_{\rm I}}(V_1), \hat{P}_{\mathbf{G}_{\rm II}}(V_1), \hat{P}_{\mathbf{G}_{\rm III}}(V_1)\right\}$, $\left\{\hat{P}_{\mathbf{G}_{\rm I}}(V_2), \hat{P}_{V_1^*}(V_2)\right\}$, $\left\{\hat{P}_{\mathbf{G}_{\rm I}}(V_3), \hat{P}_{\mathbf{G}_{\rm II}}(V_3), \hat{P}_{\mathbf{G}_{\rm III}}(V_3)\right\}$, *and* $\left\{\hat{P}_{\mathbf{G}_{\rm III}}(V_4), \hat{P}_{V_3^*}(V_4)\right\}$.

VIII. Plant condition assessment layer

The objective of this layer is to evaluate the condition of B, HT, RP, and LT based on $\hat{P}^*(V_1)$, $\hat{P}^*(V_2)$, $\hat{P}^*(V_3)$, and $\hat{P}^*(V_4)$. The calculations described below involve the plant model, introduced in (31), and the previously mentioned Jeffrey's rule (see [11]).

While $\hat{P}^*(V_1)$ can be directly used for plant condition assessment, the pmf's $\hat{P}^*(V_2)$, $\hat{P}^*(V_3)$, and $\hat{P}^*(V_4)$ cannot, since the model (31) specifies $P(V_2'|G_{\rm HT})$, $P(V_3'|G_{\rm RP})$, and $P(V_4'|G_{\rm LT})$ (with V_i', $i = 2,3,4$, defined in Subsection III-E) rather than $P(V_2|G_{\rm HT})$, $P(V_3|G_{\rm RP})$, and $P(V_4|G_{\rm LT})$. Therefore, $\hat{P}^*(V_i)$,

$i = 2, 3, 4$, must be re-calculated in terms of V_i', $i = 2, 3, 4$. This is carried out as follows: Consider, for example, $\mathbf{V}_2$. Then,

$$\begin{aligned} \hat{P}^*(V_2' = \mathrm{L}'_{\mathrm{HT}}) &= \hat{P}^*(V_2 = \mathrm{VL}_{\mathrm{HT}}) + \hat{P}^*(V_2 = \mathrm{L}_{(2)\mathrm{HT}}), \\ \hat{P}^*(V_2' = \mathrm{N}'_{\mathrm{HT}}) &= \hat{P}^*(V_2 = \mathrm{L}_{(1)\mathrm{HT}}) + \hat{P}^*(V_2 = \mathrm{N}_{\mathrm{HT}}), \end{aligned} \tag{81}$$

where the right hand side of (81) is due to the events $\{V_2 = \mathrm{VL}_{\mathrm{HT}}\}$, $\{V_2 = \mathrm{L}_{(1)\mathrm{HT}}\}$, $\{V_2 = \mathrm{L}_{(2)\mathrm{HT}}\}$, and $\{V_2 = \mathrm{N}_{\mathrm{HT}}\}$ being disjoint. Using (28)-(30), we carry out similar calculations to obtain $\hat{P}^*(V_3')$ and $\hat{P}^*(V_4')$ as well. Given these pmf's, we ascertain the condition of B, HT, RP, and LT as described below.

For the case of B, consider the following algorithm:

(a) Assign the initial pmf of G_{B} (see (25)) as

$$P_0(G_{\mathrm{B}}) = \left[\frac{1}{2},\ \frac{1}{2}\right]. \tag{82}$$

(b) Calculate the initial joint pmf of V_1 and G_{B}:

$$P_0(V_1, G_{\mathrm{B}}) = P_0(G_{\mathrm{B}})P(V_1|G_{\mathrm{B}}), \tag{83}$$

where $P(V_1|G_{\mathrm{B}})$ is given in (31).

(c) Calculate the marginal probability

$$P_0(V_1) = \sum_{G_{\mathrm{B}} \in \Sigma_{G_{\mathrm{B}}}} P_0(V_1, G_{\mathrm{B}}). \tag{84}$$

(d) Apply Jeffrey's rule (see [11]):

$$\hat{P}(V_1, G_{\mathrm{B}}) = P_0(V_1, G_{\mathrm{B}}) \frac{\hat{P}^*(V_1)}{P_0(V_1)}. \tag{85}$$

(e) Marginalize the left hand side of (85) to obtain the assessment of the condition of B:

$$\hat{P}(G_{\mathrm{B}}) = \sum_{V_1 \in \Sigma_1} \hat{P}(V_1, G_{\mathrm{B}}). \tag{86}$$

Similarly, for the other components, the pmf's $\hat{P}(G_{\mathrm{HT}})$, $\hat{P}(G_{\mathrm{RP}})$, and $\hat{P}(G_{\mathrm{LT}})$ are calculated by following steps (a)-(e) above. Thus, the overall plant assessment is expressed as a vector:

$$\hat{\vec{G}} \triangleq \left[\hat{P}(G_{\mathrm{B}}),\ \hat{P}(G_{\mathrm{HT}}),\ \hat{P}(G_{\mathrm{RP}}),\ \hat{P}(G_{\mathrm{LT}})\right]. \tag{87}$$

The pmf's in (87) comprise the output of the monitoring system that is reported to the plant operator. The next section presents performance evaluation of this monitoring system using the power plant of Figure 1.

IX. Application of Resilient Monitoring System to Power Plant

A. Parameters of power plant and monitoring system

The range and d.c. gain of process variable $\mathbf{V}_i$, $i \in \{1,2,3,4\}$, are selected as follows: Referring to (2) and (5), we assume that $V^{\mathrm{B}}_{\min} = 5$, $V^{\mathrm{B}}_{\max} = 100$, $R^{\mathrm{B}} = 50$, $\alpha_{\mathrm{L_B}}(\tilde{V}_1) = 0.005\tilde{V}_1 + 1.95$, and $\alpha_{\mathrm{N_B}}(\tilde{V}_1) = 0.0004\tilde{V}_1 + 2.18$; Similarly, for the other process variables, $V^{\mathrm{HT}}_{\min} = 5$, $V^{\mathrm{HT}}_{\max} = 25$, $R^{\mathrm{HT}}_1 = 10$, $R^{\mathrm{HT}}_2 = 15$, $R^{\mathrm{HT}}_3 = 20$, $\alpha_{\mathrm{VL_{HT}}}(\tilde{V}_2) = 0.001\tilde{V}_2 + 0.495$, $\alpha_{\mathrm{L_{(1)HT}}}(\tilde{V}_2) = 0.01\tilde{V}_2 + 0.40$, $\alpha_{\mathrm{L_{(2)HT}}}(\tilde{V}_2) = 0.005\tilde{V}_2 + 0.472$, $\alpha_{\mathrm{N_{HT}}}(\tilde{V}_2) = 0.001\tilde{V}_2 + 0.56$, $V^{\mathrm{RP}}_{\min} = 5$, $V^{\mathrm{RP}}_{\max} = 100$, $R^{\mathrm{RP}}_1 = 20$, $R^{\mathrm{RP}}_2 = 40$, $R^{\mathrm{RP}}_3 = 50$, $\alpha_{\mathrm{VL_{RP}}}(\tilde{V}_3) = 0.0016\tilde{V}_3 + 0.492$, $\alpha_{\mathrm{L_{(1)RP}}}(\tilde{V}_3) = 0.0026\tilde{V}_3 + 0.472$, $\alpha_{\mathrm{L_{(2)RP}}}(\tilde{V}_3) = \alpha_{\mathrm{L_{(1)RP}}}(\tilde{V}_3)$, $\alpha_{\mathrm{N_{RP}}}(\tilde{V}_3) = 0.0001\tilde{V}_3 + 0.6$, $V^{\mathrm{LT}}_{\min} = 0.1$, $V^{\mathrm{LT}}_{\max} = 20$, $R^{\mathrm{LT}}_1 = 3$, $R^{\mathrm{LT}}_2 = 6$, $R^{\mathrm{LT}}_3 = 9$, $R^{\mathrm{LT}}_4 = 11$, $R^{\mathrm{LT}}_5 = 13$, $R^{\mathrm{LT}}_6 = 15$, $R^{\mathrm{LT}}_7 = 17$, $\alpha_{\mathrm{VL_{(1)LT}}}(\tilde{V}_4) = 0.0014\tilde{V}_4 + 0.4$, $\alpha_{\mathrm{VL_{(2)LT}}}(\tilde{V}_4) = 0.006\tilde{V}_4 + 0.38$, $\alpha_{\mathrm{L_{(1)LT}}}(\tilde{V}_4) = 0.001\tilde{V}_4 + 0.416$, $\alpha_{\mathrm{L_{(2)LT}}}(\tilde{V}_4) = 0.01\tilde{V}_4 + 0.332$, $\alpha_{\mathrm{M_{(1)LT}}}(\tilde{V}_4) = 0.002\tilde{V}_4 + 0.426$, $\alpha_{\mathrm{M_{(2)LT}}}(\tilde{V}_4) = 0.01\tilde{V}_4 + 0.3$, $\alpha_{\mathrm{N_{LT}}}(\tilde{V}_4) = 0.002\tilde{V}_4 + 0.441$, and $\alpha_{\mathrm{H_{LT}}}(\tilde{V}_4) = 0.008\tilde{V}_4 + 0.35$. Note that the temperature values and d.c. gains introduced above are just illustrative and can be scaled as needed. Further, we assume that $\tilde{V}_i$, $i \in \{1,2,3,4\}$, is a Gaussian random variable, whose pdf is specified by $\mathcal{N}\left(\mu_{\tilde{V}_i}, \sigma_{\tilde{V}_i}\right)$, with the standard deviation $\sigma_{\tilde{V}_i}$ being sufficiently small so that realizations of $\tilde{V}_i$ outside of the corresponding interval (e.g., $[V^{\mathrm{B}}_{\min}, V^{\mathrm{B}}_{\max}]$) can be ignored.

The random variable $\tilde{S}_{ij}$, $i \in \{1,2,3,4\}$, $j \in \{1,2\}$, which characterizes the measurements of sensor $\mathbf{S}_{ij}$, is assumed to be distributed according to $\mathcal{N}\left(\mu_{\tilde{S}_{ij}}, \sigma_{\tilde{S}_{ij}}\right)$. The sampling period of all sensors is taken as 0.01sec.

Regarding the probing signals associated with the DQ layer, their magnitudes are assigned as follows: $A^{V_1}_0 = 2$ and $A^{V_2}_0 = A^{V_3}_0 = A^{V_4}_0 = 0.7$. The DQ parameter ϵ, involved in (42), is selected as 0.02.

Regarding the process variable pmf evaluation procedure, the parameter ϵ_h, involved in (46), is assigned as 0.01. The stopping rule of (46) is given by

$$|h_{\sigma_1}(n+1) - h_{\sigma_1}(n)| < 10^{-4}. \tag{88}$$

The level of rationality N of the rational controllers is chosen to be 2. The parameters η and $T_{\max}$, involved in (58), are taken as 0.04 and 1, respectively.

Regarding the plant model, the conditional pmf's involved in (31) are assumed as follows:

- $P(V_1 = \mathrm{N_B}|G_{\mathrm{B}} = \mathrm{N_{G_B}}) = P(V_1 = \mathrm{L_B}|G_{\mathrm{B}} = \mathrm{A_{G_B}}) = 0.95$ and other probabilities are 0.05;
- $P(V'_2 = \mathrm{N'_{HT}}|G_{\mathrm{HT}} = \mathrm{N_{G_{HT}}}) = P(V'_2 = \mathrm{L'_{HT}}|G_{\mathrm{HT}} = \mathrm{A_{G_{HT}}}) = 0.90$ and other probabilities are 0.1;
- $P(V'_3 = \mathrm{N'_{RP}}|G_{\mathrm{RP}} = \mathrm{N_{G_{RP}}}) = 0.88$, $P(V'_3 = \mathrm{L'_{RP}}|G_{\mathrm{RP}} = \mathrm{A_{G_{RP}}}) = 0.91$, $P(V'_3 = \mathrm{L'_{RP}}|G_{\mathrm{RP}} = \mathrm{N_{G_{RP}}}) = 0.12$, and $P(V'_3 = \mathrm{N'_{RP}}|G_{\mathrm{RP}} = \mathrm{A_{G_{RP}}}) = 0.09$;

- $P(V_4' = \mathrm{N}'_{\mathrm{LT}}|G_{\mathrm{LT}} = \mathrm{N}_{\mathrm{G_{LT}}}) = 0.91$, $P(V_4' = \mathrm{H}'_{\mathrm{LT}}|G_{\mathrm{LT}} = \mathrm{A}_{\mathrm{G_{LT}}}) = 0.92$, $P(V_4' = \mathrm{H}'_{\mathrm{LT}}|G_{\mathrm{LT}} = \mathrm{N}_{\mathrm{G_{LT}}}) = 0.09$, and $P(V_4' = \mathrm{N}'_{\mathrm{LT}}|G_{\mathrm{LT}} = \mathrm{A}_{\mathrm{G_{LT}}}) = 0.08$.

In all scenarios considered, the plant condition is evaluated and expressed as

$$\hat{\vec{G}} \triangleq \left[\hat{P}(G_{\mathrm{B}}),\ \hat{P}(G_{\mathrm{HT}}),\ \hat{P}(G_{\mathrm{RP}}),\ \hat{P}(G_{\mathrm{LT}})\right], \tag{89}$$

wherein the pmf's are represented by the row vectors $\hat{P}(G_{\mathrm{B}}) = [\hat{P}(G_{\mathrm{B}} = \mathrm{N}_{\mathrm{G_B}}), \hat{P}(G_{\mathrm{B}} = \mathrm{A}_{\mathrm{G_B}})]$, $\hat{P}(G_{\mathrm{HT}}) = [\hat{P}(G_{\mathrm{HT}} = \mathrm{N}_{\mathrm{G_{HT}}}), \hat{P}(G_{\mathrm{HT}} = \mathrm{A}_{\mathrm{G_{HT}}})]$, $\hat{P}(G_{\mathrm{RP}}) = [\hat{P}(G_{\mathrm{RP}} = \mathrm{N}_{\mathrm{G_{RP}}}), \hat{P}(G_{\mathrm{RP}} = \mathrm{A}_{\mathrm{G_{RP}}})]$, and $\hat{P}(G_{\mathrm{LT}}) = [\hat{P}(G_{\mathrm{LT}} = \mathrm{N}_{\mathrm{G_{LT}}}), \hat{P}(G_{\mathrm{LT}} = \mathrm{A}_{\mathrm{G_{LT}}})]$.

B. Performance evaluation

In all figures of this subsection, states of the sensor subnetworks (see (23)) are indexed as follows: For $k \in \{\mathrm{I}, \mathrm{II}, \mathrm{III}\}$,

$$\begin{aligned}
&1:\ (1111)_k,\ 2:\ (1110)_k,\ 3:\ (1101)_k,\ 4:\ (1011)_k,\\
&5:\ (1010)_k,\ 6:\ (1001)_k,\ 7:\ (0111)_k,\ 8:\ (0110)_k,\\
&9:\ (0101)_k,\ 10:\ (1000)_k,\ 11:\ (0100)_k,\ 12:\ (1100)_k,\\
&13:\ (0010)_k,\ 14:\ (0001)_k,\ 15:\ (0011)_k.
\end{aligned} \tag{90}$$

Regarding the non-resilient system to which the resilient one is compared, it so happens that in some situations, Dempster-Shafer rule cannot be applied to combine two candidate pmf's. For example, suppose $\hat{P}_{\mathrm{nr}}^{S_{11}}(V_1) = [1, 0]$ and $\hat{P}_{\mathrm{nr}}^{S_{12}}(V_1) = [0, 1]$; Clearly, combining these two pmf's using Dempster-Shafer rule is not possible since it results in $\left[\frac{0}{0},\ \frac{0}{0}\right]$. Therefore, to compute $\hat{P}_{\mathrm{nr}}^{S_{11}S_{12}}(V_1)$, we utilize the following rule:

$$\hat{P}_{\mathrm{nr}}^{S_{11}S_{12}}(V_1) = \frac{1}{2}\left(\hat{P}_{\mathrm{nr}}^{S_{11}}(V_1) + \hat{P}_{\mathrm{nr}}^{S_{12}}(V_1)\right). \tag{91}$$

The performance of the monitoring system designed in this paper is evaluated in the framework of the following three scenarios:

Scenario 1: All power plant components are operating normally, with $\mu_{\tilde{V}_1} = 80$, $\mu_{\tilde{V}_2} = 21$, $\mu_{\tilde{V}_3} = 60$, $\mu_{\tilde{V}_4} = 15.3$, and $\sigma_{\tilde{V}_1} = \sigma_{\tilde{V}_2} = \sigma_{\tilde{V}_3} = \sigma_{\tilde{V}_4} = 0.01$. Sensors $\mathbf{S}_{11}$, $\mathbf{S}_{21}$, $\mathbf{S}_{31}$, and $\mathbf{S}_{41}$ are captured and forced to indicate malfunctioning of LT and fractures in the insulations of B, HT, and RP. The sensor distributions are characterized by $\mu_{\tilde{S}_{11}} = 28.5$, $\sigma_{\tilde{S}_{11}} = 0.2$, $\mu_{\tilde{S}_{12}} = 79$, $\sigma_{\tilde{S}_{12}} = 0.21$, $\mu_{\tilde{S}_{21}} = 8.5$, $\sigma_{\tilde{S}_{21}} = 0.12$, $\mu_{\tilde{S}_{22}} = 22.3$, $\sigma_{\tilde{S}_{22}} = 0.1$, $\mu_{\tilde{S}_{31}} = 16.2$, $\sigma_{\tilde{S}_{31}} = 0.1$, $\mu_{\tilde{S}_{32}} = 61$, $\sigma_{\tilde{S}_{32}} = 0.11$, $\mu_{\tilde{S}_{41}} = 5.1$, $\sigma_{\tilde{S}_{41}} = 0.1$, $\mu_{\tilde{S}_{42}} = 15.3$, and $\sigma_{\tilde{S}_{42}} = 0.1$. Based on these data, sensor DQ's are evaluated as $DQ_{11} = 0.015$, $DQ_{12} = 0.99$, $DQ_{21} \approx 0$, $DQ_{22} = 0.96$, $DQ_{31} \approx 0$, $DQ_{32} = 0.96$, $DQ_{41} \approx 0$, and $DQ_{42} = 0.95$.

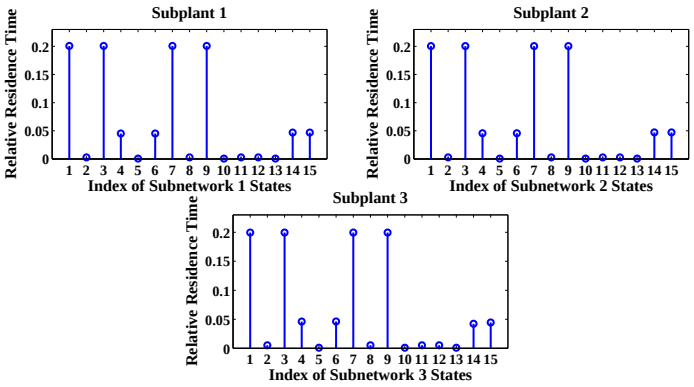


Fig. 5: Relative residence time for Scenario 1

The performance of the monitoring system is illustrated in Figure 5. It indicates that the residence time in states, which take into account measurements of only captured sensors is small. The plant component pmf's are computed as $\hat{P}(G_{\text{B}}) = [0.95, 0.05]$, $\hat{P}(G_{\text{HT}}) = [0.89, 0.11]$, $\hat{P}(G_{\text{RP}}) = [0.91, 0.09]$, and $\hat{P}(G_{\text{LT}}) = [0.91, 0.09]$, which indicate accurately that all components are operating normally.

Given that four sensors indicate normal, while the other four indicate damage/malfunctioning, the non-resilient system evaluates the above pmf's to be approximately uniform ($[0.5, 0.5]$). This leads to the measure of resiliency $\overrightarrow{MR} = [0.93, 0.83, 0.85, 0.87]$, testifying to the efficacy of the resilient monitoring system.

Scenario 2: LT is malfunctioning and other components operate normally, with $\mu_{\tilde{V}_1} = 80$, $\mu_{\tilde{V}_2} = 24.5$, $\mu_{\tilde{V}_3} = 75$, $\mu_{\tilde{V}_4} = 18.4$, and $\sigma_{\tilde{V}_1} = \sigma_{\tilde{V}_2} = \sigma_{\tilde{V}_3} = \sigma_{\tilde{V}_4} = 0.01$. Sensors $\mathbf{S}_{11}$, $\mathbf{S}_{12}$, $\mathbf{S}_{21}$, and $\mathbf{S}_{22}$ are captured and forced to indicate that B and HT are damaged. The sensor distributions are assumed to be characterized by $\mu_{\tilde{S}_{11}} = 31$, $\sigma_{\tilde{S}_{11}} = 0.2$, $\mu_{\tilde{S}_{12}} = 31$, $\sigma_{\tilde{S}_{12}} = 0.21$, $\mu_{\tilde{S}_{21}} = 15.1$, $\sigma_{\tilde{S}_{21}} = 0.12$, $\mu_{\tilde{S}_{22}} = 15.2$, $\sigma_{\tilde{S}_{22}} = 0.1$, $\mu_{\tilde{S}_{31}} = 74$, $\sigma_{\tilde{S}_{31}} = 0.11$, $\mu_{\tilde{S}_{32}} = 74.1$, $\sigma_{\tilde{S}_{32}} = 0.1$, $\mu_{\tilde{S}_{41}} = 18.1$, $\sigma_{\tilde{S}_{41}} = 0.1$, $\mu_{\tilde{S}_{42}} = 18.6$, $\sigma_{\tilde{S}_{42}} = 0.1$. Based on these data, sensor DQ's are evaluated as $DQ_{11} = DQ_{12} = 0.015$, $DQ_{21} = DQ_{22} = 0.006$, $DQ_{31} = DQ_{32} = 1$, and $DQ_{41} = DQ_{42} = 0.99$.

The performance of the monitoring system is illustrated in Figure 6. The residence time in all states of subnetwork X_{I} is the same, since all its sensors are captured. Despite the lack of reliable information from the sensors of B, the monitoring system calculates $\hat{P}(G_{\text{B}}) = [0.95, 0.05]$. This is due to the inter-subplant inferences (see Subsection VII). Further, the pmf $\hat{P}(G_{\text{HT}})$ is calculated as $[0.49, 0.51]$, since it is impossible to determine the condition of HT using sensors of other components. Finally, regarding RP and LT, the following pmf's are computed: $\hat{P}(G_{\text{RP}}) = [0.91, 0.09]$ and $\hat{P}(G_{\text{LT}}) = [0.09, 0.91]$. The above pmf's

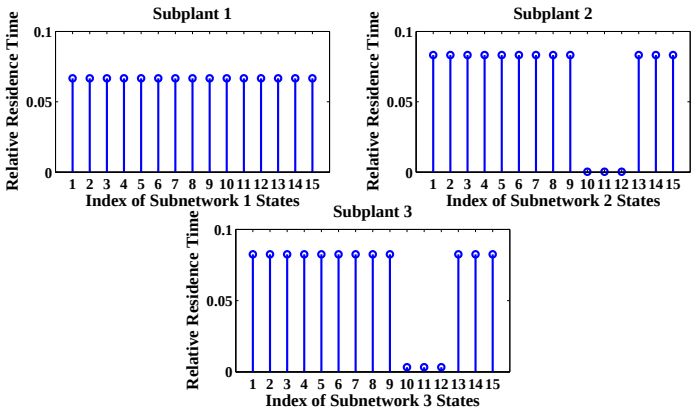


Fig. 6: Relative residence time for Scenario 2

indicate accurately that B and RP operate normally, while LT is malfunctioning.

The non-resilient system evaluates $\hat{P}_{\text{nr}}(G_{\text{B}}) = [0.05, 0.95]$ and $\hat{P}_{\text{nr}}(G_{\text{HT}}) = [0.1, 0.9]$, which erroneously indicate that the insulations of B and HT are damaged. This leads to the measure of resiliency $\overrightarrow{MR} = [0.98, 0.7, -, -]$, where '$-$' is used to indicate that none of the sensors of a particular component are captured. Once again, these results demonstrate the efficacy of the monitoring system presented in this work.

Scenario 3: All power plant components function normally, with $\mu_{\tilde{V}_1} = 80$, $\mu_{\tilde{V}_2} = 24.5$, $\mu_{\tilde{V}_3} = 75$, $\mu_{\tilde{V}_4} = 16$, and $\sigma_{\tilde{V}_1} = \sigma_{\tilde{V}_2} = \sigma_{\tilde{V}_3} = \sigma_{\tilde{V}_4} = 0.01$. All sensors are captured and indicate malfunctioning of all power plant components. The sensor distributions are characterized by $\mu_{\tilde{S}_{11}} = 31$, $\sigma_{\tilde{S}_{11}} = 0.2$, $\mu_{\tilde{S}_{12}} = 31$, $\sigma_{\tilde{S}_{12}} = 0.21$, $\mu_{\tilde{S}_{21}} = 15.1$, $\sigma_{\tilde{S}_{21}} = 0.12$, $\mu_{\tilde{S}_{22}} = 15.2$, $\sigma_{\tilde{S}_{22}} = 0.1$, $\mu_{\tilde{S}_{31}} = 5$, $\sigma_{\tilde{S}_{31}} = 0.11$, $\mu_{\tilde{S}_{32}} = 5.1$, $\sigma_{\tilde{S}_{32}} = 0.1$, $\mu_{\tilde{S}_{41}} = 4.1$, $\sigma_{\tilde{S}_{41}} = 0.1$, $\mu_{\tilde{S}_{42}} = 4.2$, $\sigma_{\tilde{S}_{42}} = 0.2$. Based on these data, all DQ's are computed to be close to 0.

The performance of the monitoring system is illustrated in Figure 7. As expected, the residence time in all subnetwork states is the same. The pmf's $\hat{P}(G_{\text{B}})$, $\hat{P}(G_{\text{HT}})$, $\hat{P}(G_{\text{RP}})$, and $\hat{P}(G_{\text{LT}})$ are computed to be uniform ($[0.5, 0.5]$), which is reasonable since no trustworthy information about any of the components is available.

The non-resilient system evaluates $\hat{P}_{\text{nr}}(G_{\text{B}}) = [0.05, 0.95]$, $\hat{P}_{\text{nr}}(G_{\text{HT}}) = [0.1, 0.9]$, $\hat{P}_{\text{nr}}(G_{\text{RP}}) = [0.12, 0.88]$, and $\hat{P}_{\text{nr}}(G_{\text{LT}}) = [0.09, 0.91]$, which erroneously indicate that all components are damaged. The measure of resiliency is computed as $\overrightarrow{MR} = [0.76, 0.7, 0.69, 0.72]$, which testifies to the efficacy of the monitoring system presented in this work.

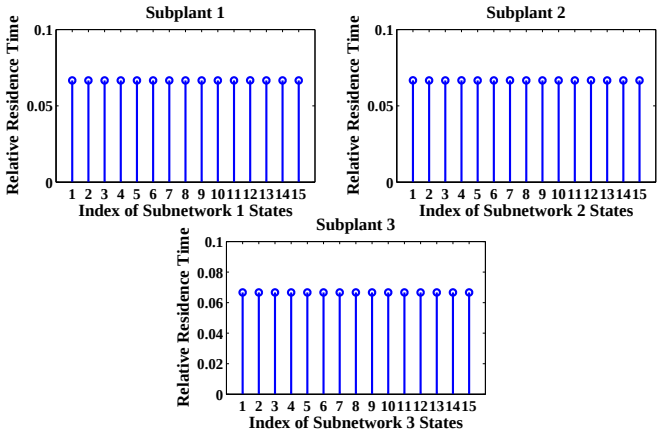


Fig. 7: Relative residence time for Scenario 3

X. Conclusions and Future work

This paper developed an approach to resilient monitoring of power plants based on a five-layer architecture, which includes an overlapping decomposition with intra- and inter-subplant inferences. The calculations used at each layer are described and justified. Numerical performance evaluation of the system developed testifies to its efficacy.

Numerous problems in this area still remain open. The main one, in our opinion, is the relatively low rate of plant condition assessment. In the current implementation, plant assessment results are obtained every 180sec. At least an order of magnitude improvement would be desirable. This may be accomplished using more efficient rational controllers for adaptation at the process variable assessment layer and more expedient data quality assessment algorithms. Other open problems include:

- Development of the overlapping decomposition approach for plants with cyclic (rather than tree-type) influence diagrams among the plant components.
- Improving models of process variables, plant, and the attacker by making them more general and practical. For example, attackers other than mean-based should be introduced and analyzed.
- Practical application of the developed resilient monitoring systems is a challenging task for future research.

Solution of these problems will lead to a relatively complete and useful theory of resilient monitoring systems for plants under cyber-physical attacks.

APPENDIX

The proof of Theorem 7.1 is based on the following five lemmas.

Lemma A.1: Let $\hat{P}_{x_{\rm I}}(V_1)$ be the pmf of V_1 calculated as mentioned in Subsection VII-B and $\hat{P}_{x_{\rm I}}(V_2)$ be calculated using total probability formula:

$$\hat{P}_{x_{\rm I}}(V_2) = \sum_{\sigma_1 \in \Sigma_1} P(V_2|V_1 = \sigma_1)\hat{P}_{x_{\rm I}}(V_1 = \sigma_1), \tag{92}$$

where $P(V_2|V_1)$ and Σ_1 are specified by (17) and (67), respectively. Then,

$$I\left\{\hat{P}_{x_{\rm I}}(V_2)\right\} = \frac{I\left\{\hat{P}_{x_{\rm I}}(V_1)\right\} + 1}{2},\ x_{\rm I} \in X_{\rm I}, \tag{93}$$

where $I\{.\}$ is the entropy.

Proof: Due to the structure of $P(V_2|V_1)$, equation (92) can be expressed as $\hat{P}_{x_{\rm I}}(V_2 = {\rm N}_2) = \hat{P}_{x_{\rm I}}(V_2 = {\rm A}_{21}) = \frac{1}{2}\hat{P}_{x_{\rm I}}(V_1 = {\rm N}_1)$ and $\hat{P}_{x_{\rm I}}(V_2 = {\rm A}_{22}) = \hat{P}_{x_{\rm I}}(V_2 = {\rm A}_{23}) = \frac{1}{2}\hat{P}_{x_{\rm I}}(V_1 = {\rm A}_1)$. Consequently, the entropy of $\hat{P}_{x_{\rm I}}(V_1)$ can be evaluated as

$$I\left\{\hat{P}_{x_{\rm I}}(V_1)\right\} = -\sum_{\sigma_2 \in \Sigma_2} \hat{P}_{x_{\rm I}}(V_2 = \sigma_2)\log_2 \hat{P}_{x_{\rm I}}(V_2 = \sigma_2) - 1, \tag{94}$$

where Σ_2 is defined in (67). Then, applying the change of base formula, $\log_a x = \frac{\log_b x}{\log_b a}$, the right hand side (RHS) of (94) becomes $I\left\{\hat{P}_{x_{\rm I}}(V_2)\right\}\log_2 4 - 1 = 2I\left\{\hat{P}_{x_{\rm I}}(V_2)\right\} - 1$. □

Lemma A.2: Let $\hat{P}_x(V_1, V_2)$ be the joint pmf of V_1 and V_2 calculated as mentioned in Subsection VII-B and $\hat{P}_x(V_2)$ be the pmf of V_2 obtained by marginalizing $\hat{P}_x(V_1, V_2)$. Then,

$$I\left\{\hat{P}_x(V_2)\right\} = \frac{3}{2}I\left\{\hat{P}_x(V_1, V_2)\right\},\ x \in X, \tag{95}$$

where $I\{.\}$ is, as before, the entropy.

The proof of this lemma is similar to that of Lemma A.1. □

Since $\hat{P}_{x^*}(V_1, V_2)$ and $\hat{P}_{x_{\rm I}^*}(V_1)$ possess the smallest entropy in the state space of their respective networks (see (72) and (76)), Lemmas A.1 and A.2 indicate that the global and local rational controllers adapt in such a manner that the entropy of the pmf of V_2 is minimized.

The following three lemmas provide technical conditions for the proof of Theorem 7.1:

Lemma A.3: Assume that $\hat{P}_{x_{\rm I}}(V_2)$ and $\hat{P}_{x_{\rm II}}(V_2)$, $x_{\rm I} \in X_{\rm I}$, $x_{\rm II} \in X_{\rm II}$, are DS-monotonic and $\hat{P}_{(x_{\rm I}, x_{\rm II})}(V_2)$ is their concatenation using Dempster-Shafer rule. Then,

$$\frac{dI\left\{\hat{P}_{(x_{\rm I}, x_{\rm II})}(V_2)\right\}}{dI\left\{\hat{P}_{x_{\rm I}}(V_2)\right\}} > 0 \text{ and } \frac{dI\left\{\hat{P}_{(x_{\rm I}, x_{\rm II})}(V_2)\right\}}{dI\left\{\hat{P}_{x_{\rm II}}(V_2)\right\}} > 0. \tag{96}$$

Proof: Introduce notations: $\hat{P}_{x_{\rm I}}(V_2) = [p_1, p_2, p_3, p_4]$ and $\hat{P}_{x_{\rm II}}(V_2) = [q_1, q_2, q_3, q_4]$. Without loss of generality, assume that $\max\{p_1, p_2, p_3, p_4\} = p_1$ and $\max\{q_1, q_2, q_3, q_4\} = q_1$. The entropy of $\hat{P}_{x_{\rm I}}(V_2)$ is

$$I\left\{\hat{P}_{x_{\rm I}}(V_2)\right\} = -\sum_{i=1}^{4} p_i \log_4 p_i. \tag{97}$$

The differential of $I\left\{\hat{P}_{x_{\rm I}}(V_2)\right\}$ is

$$dI\left\{\hat{P}_{x_{\rm I}}(V_2)\right\} = \sum_{i=1}^{4} \frac{\partial}{\partial p_i} I\left\{\hat{P}_{x_{\rm I}}(V_2)\right\} dp_i, \tag{98}$$

where, due to the constraint, $p_1 + p_2 + p_3 + p_4 = 1$,

$$\sum_{i=1}^{4} dp_i = 0. \tag{99}$$

Using (99), equation (98) can be re-written as follows:

$$dI\left\{\hat{P}_{x_{\rm I}}(V_2)\right\} = \sum_{i=1}^{3} \left[\frac{\partial}{\partial p_i} - \frac{\partial}{\partial p_4}\right] I\left\{\hat{P}_{x_{\rm I}}(V_2)\right\} dp_i. \tag{100}$$

Since, as it follows from (97), the partial derivative, $\frac{\partial}{\partial p_i} I\left\{\hat{P}_{x_{\rm I}}(V_2)\right\} = -1 - \log_4 p_i$, $i = 1, 2, 3, 4$. From (100), we obtain the following differential for the denominator of the first expression of (96):

$$dI\left\{\hat{P}_{x_{\rm I}}(V_2)\right\} = \sum_{i=1}^{3} \left[\log_4 \frac{p_4}{p_i}\right] dp_i. \tag{101}$$

In a similar manner, it can be shown that the numerator of this expression is given by:

$$\begin{aligned} &dI\left\{\hat{P}_{(x_{\rm I}, x_{\rm II})}(V_2)\right\} = \\ &\sum_{i=1}^{3} \left[\left(\frac{Dq_4 - p_4 q_4^2}{D^2}\right)\left(1 + \log_4 \frac{p_4 q_4}{D}\right) - \left(\frac{Dq_i - p_i q_i^2}{D^2}\right)\left(1 + \log_4 \frac{p_i q_i}{D}\right)\right] dp_i, \end{aligned} \tag{102}$$

where

$$D = \sum_{i=1}^{4} p_i q_i. \tag{103}$$

Suppose, $dp_1 > 0$ and $dp_2 = dp_3 = 0$. Then, the RHS of (101) becomes $\left[\log_4 \frac{p_4}{p_1}\right] dp_1$, implying that $dI\left\{\hat{P}_{x_{\rm I}}(V_2)\right\} < 0$. Using (102), it can also be shown that $dI\left\{\hat{P}_{(x_{\rm I}, x_{\rm II})}(V_2)\right\} < 0$. Moreover, in all other situations where $dI\left\{\hat{P}_{x_{\rm I}}(V_2)\right\}$ is less than zero (for instance, when $dp_1 > 0$, $dp_2 > 0$, $dp_3 = 0$, $p_2 > p_4$, and $q_2 > q_4$), it can be shown that $dI\left\{\hat{P}_{(x_{\rm I}, x_{\rm II})}(V_2)\right\}$ is less than zero as well. These arguments imply that $\frac{dI\left\{\hat{P}_{(x_{\rm I}, x_{\rm II})}(V_2)\right\}}{dI\left\{\hat{P}_{x_{\rm I}}(V_2)\right\}} > 0$.

The second expression in (96) can be proved in a similar manner. □

Lemma A.4: Let x_k, $k = \mathrm{I}, \mathrm{II}$, be a state in X_k, with all active sensors being captured. Let the pmf $\hat{P}_{x_k}(V_2)$ be computed as mentioned in Subsection VII-B. Finally, let $\epsilon << 1$ be the parameter involved in the data quality exponent (42). Then, $I\left\{\hat{P}_{x_k}(V_2)\right\} = 1 - \delta(\epsilon)$, $\delta(\epsilon) \to 0$ as $\epsilon \to 0$.

Proof: We prove this lemma for the state $x_{\mathrm{I}} \in X_{\mathrm{I}}$. The proof for $x_{\mathrm{II}} \in X_{\mathrm{II}}$ is similar.

Suppose, $x_{\mathrm{I}} = (10)_{\mathrm{I}}$. Let the pmf of the corresponding sensor, namely, $\mathbf{S}_{11}$, be $P(S_{11}) = \left[P_{\mathrm{L_B}}^{S_{11}}, P_{\mathrm{N_B}}^{S_{11}}\right]$. Since this sensor is captured, its data quality is assigned as $DQ_{11} = \epsilon$ based on the d.c. gain model (68) and the procedure described in Section IV. In this situation, the pmf $\hat{P}_{x_{\mathrm{I}}}(V_1)$, calculated using the h-procedure, is as follows (see (48)):

$$\hat{P}_{x_{\mathrm{I}}}(V_1) = \left[P_{\mathrm{L_B}}^{S_{11}}\epsilon + \frac{1-\epsilon}{2}, P_{\mathrm{N_B}}^{S_{11}}\epsilon + \frac{1-\epsilon}{2}\right]. \tag{104}$$

The entropy of $\hat{P}_{x_{\mathrm{I}}}(V_1)$ is

$$I\left\{\hat{P}_{x_{\mathrm{I}}}(V_1)\right\} = -a_1(\epsilon)\log_2 a_1(\epsilon) - a_2(\epsilon)\log_2 a_2(\epsilon), \tag{105}$$

where

$$\begin{aligned} a_1(\epsilon) &:= P_{\mathrm{L_B}}^{S_{11}}\epsilon + \tfrac{1-\epsilon}{2}, \\ a_2(\epsilon) &:= P_{\mathrm{N_B}}^{S_{11}}\epsilon + \tfrac{1-\epsilon}{2}. \end{aligned} \tag{106}$$

Taking into account (105) and using Lemma A.1, we have

$$I\left\{\hat{P}_{x_{\mathrm{I}}}(V_2)\right\} = \frac{1}{2}\left[1 - a_1(\epsilon)\log_2 a_1(\epsilon) - a_2(\epsilon)\log_2 a_2(\epsilon)\right]. \tag{107}$$

As it follows from (106), the expression $-a_1(\epsilon)\log_2 a_1(\epsilon) - a_2(\epsilon)\log_2 a_2(\epsilon) \to 1$ as $\epsilon \to 0$.

Differentiating both sides of (107) with respect to ϵ, we obtain

$$\frac{d}{d\epsilon} I\left\{\hat{P}_{x_{\mathrm{I}}}(V_2)\right\} = \left(P_{\mathrm{L_B}}^{S_{11}} - \frac{1}{2}\right)\log_2\frac{a_2(\epsilon)}{a_1(\epsilon)}. \tag{108}$$

It follows from (106) and (108) that $\frac{d}{d\epsilon} I\left\{\hat{P}_{x_{\mathrm{I}}}(V_2)\right\} \leq 0$, and the equality is attained when $P_{\mathrm{L_B}}^{S_{11}} = \frac{1}{2}$. In other words, the entropy of $\hat{P}_{x_{\mathrm{I}}}(V_2)$ is a decreasing function of ϵ.

Equation (107) can be re-expressed as

$$I\left\{\hat{P}_{x_{\mathrm{I}}}(V_2)\right\} = 1 + \frac{-1 - a_1(\epsilon)\log_2 a_1(\epsilon) - a_2(\epsilon)\log_2 a_2(\epsilon)}{2}. \tag{109}$$

Denote $\frac{-1-a_1(\epsilon)\log_2 a_1(\epsilon) - a_2(\epsilon)\log_2 a_2(\epsilon)}{2}$ as $\delta(\epsilon)$. It is seen that $\delta(\epsilon) \to 0$ as $\epsilon \to 0$.

Similarly, if x_{I} is either $(01)_{\mathrm{I}}$ or $(11)_{\mathrm{I}}$, it can be shown that $I\left\{\hat{P}_{x_{\mathrm{I}}}(V_2)\right\}$ is equal to $1 - \delta(\epsilon)$, where $\delta(\epsilon) \to 0$ as $\epsilon \to 0$. □

Lemma A.5: Let the pmf's $\hat{P}_{x_{\rm I}^*}(V_2)$ and $\hat{P}_{x_{\rm II}^*}(V_2)$ be DS-monotonic. Further, let $x_{\rm I}$ and $x_{\rm II}$ be states in $X_{\rm I}$ and $X_{\rm II}$, respectively, with their corresponding active sensors being captured. Then, the following is satisfied:

$$I\left\{\hat{P}_{(x_{\rm I}^*,x_{\rm II}^*)}(V_2)\right\} < \min\left[I\left\{\hat{P}_{(x_{\rm I}^*,x_{\rm II})}(V_2)\right\}, I\left\{\hat{P}_{(x_{\rm I},x_{\rm II}^*)}(V_2)\right\}, I\left\{\hat{P}_{(x_{\rm I},x_{\rm II})}(V_2)\right\}\right]. \tag{110}$$

Proof: If the sensors corresponding to $x_{\rm II}$ are captured, it follows from Lemma A.4 that $\hat{P}_{x_{\rm II}}(V_2) = \left[\frac{1}{4}+\delta_1(\epsilon), \frac{1}{4}+\delta_2(\epsilon), \frac{1}{4}+\delta_3(\epsilon), \frac{1}{4}+\delta_4(\epsilon)\right]$, where $\delta_i(\epsilon) \to 0$ as $\epsilon \to 0$, $i = 1, 2, 3, 4$. Consequently, the entropy of the pmf $\hat{P}_{(x_{\rm I}^*,x_{\rm II})}(V_2)$, which is obtained by concatenating $\hat{P}_{x_{\rm I}^*}(V_2)$ and $\hat{P}_{x_{\rm II}}(V_2)$ using, as before, Dempster-Shafer rule, can be shown to have the following property: With ϵ being the parameter involved in the DQ exponent (42),

$$I\left\{\hat{P}_{(x_{\rm I}^*,x_{\rm II})}(V_2)\right\} = [1-\delta_1(\epsilon)]\, I\left\{\hat{P}_{x_{\rm I}^*}(V_2)\right\} + \delta_2(\epsilon), \tag{111}$$

where $\delta_i(\epsilon) \to 0$ as $\epsilon \to 0$, $i = 1, 2$. Using (111), the difference between the entropies, $I\left\{\hat{P}_{(x_{\rm I}^*,x_{\rm II})}(V_2)\right\}$ and $I\left\{\hat{P}_{(x_{\rm I}^*,x_{\rm II}^*)}(V_2)\right\}$, can be expressed as

$$I\left\{\hat{P}_{(x_{\rm I}^*,x_{\rm II})}(V_2)\right\} - I\left\{\hat{P}_{(x_{\rm I}^*,x_{\rm II}^*)}(V_2)\right\} = [1-\delta_1(\epsilon)]\, I\left\{\hat{P}_{x_{\rm I}^*}(V_2)\right\} - I\left\{\hat{P}_{(x_{\rm I}^*,x_{\rm II}^*)}(V_2)\right\} + \delta_2(\epsilon). \tag{112}$$

Since the pmf's $\hat{P}_{x_{\rm I}^*}(V_2)$ and $\hat{P}_{x_{\rm II}^*}(V_2)$ are DS-monotonic, and ϵ is chosen arbitrarily small, the RHS of (112) is greater than zero. This implies that $I\left\{\hat{P}_{(x_{\rm I}^*,x_{\rm II}^*)}(V_2)\right\} < I\left\{\hat{P}_{(x_{\rm I}^*,x_{\rm II})}(V_2)\right\}$.

Using the above approach and Lemma A.4, it can also be shown that $I\left\{\hat{P}_{(x_{\rm I}^*,x_{\rm II}^*)}(V_2)\right\} < I\left\{\hat{P}_{(x_1,x_{\rm II}^*)}(V_2)\right\}$ and $I\left\{\hat{P}_{(x_{\rm I}^*,x_{\rm II}^*)}(V_2)\right\} < I\left\{\hat{P}_{(x_1,x_2)}(V_2)\right\}$. □

Proof of Theorem 7.1: The pmf of V_2 is evaluated using DPVA-I and CPVA as follows:

DPVA-I: Recall that $\hat{P}_{x_{\rm I}^*}(V_2)$ and $\hat{P}_{x_{\rm II}^*}(V_2)$ are DS-monotonic, due to which the pmf $\hat{P}_{(x_{\rm I}^*,x_{\rm II}^*)}(V_2)$ possesses the smallest entropy among the three of them. Thus, $\hat{P}_{(x_{\rm I}^*,x_{\rm II}^*)}(V_2)$ is reported as the pmf of V_2.

CPVA: Since $x_{\rm I}^*$ minimizes the penalty function $\Phi_{\rm I}(x_{\rm I})$, it can be shown using Lemma A.1 that $I\left\{\hat{P}_{(x_{\rm I}^*,(00))}(V_2)\right\} < I\left\{\hat{P}_{(x_1,(00))}(V_2)\right\}$, $\forall x_{\rm I} \in X_{\rm I}$. Similarly, it can also be shown that $I\left\{\hat{P}_{((00),x_{\rm II}^*)}(V_2)\right\} < I\left\{\hat{P}_{((00),x_{\rm II})}(V_2)\right\}$, $\forall x_{\rm II} \in X_{\rm II}$. Further, since the pmf's $\hat{P}_{x_{\rm I}^*}(V_2)$ and $\hat{P}_{x_{\rm II}^*}(V_2)$ are DS-monotonic, it is clear that

$$\begin{aligned} I\left\{\hat{P}_{(x_{\rm I}^*,x_{\rm II}^*)}(V_2)\right\} &< I\left\{\hat{P}_{(x_{\rm I},(00))}(V_2)\right\},\ \forall x_{\rm I} \in X_{\rm I}, \\ I\left\{\hat{P}_{(x_{\rm I}^*,x_{\rm II}^*)}(V_2)\right\} &< I\left\{\hat{P}_{((00),x_{\rm II})}(V_2)\right\},\ \forall x_{\rm II} \in X_{\rm II}. \end{aligned} \tag{113}$$

From (113) and Lemma A.2, it is obvious that the global optimum, x^*, belongs to the set $\{X_{\rm I} \times X_{\rm II}\}$.

Now, we investigate whether x^* is, indeed, $(x_{\rm I}^*, x_{\rm II}^*)$ in the following two cases:

- Case 1: For arbitrary $x_{\rm I}$ and $x_{\rm II}$, let the pmf's $\hat{P}_{x_{\rm I}}(V_2)$ and $\hat{P}_{x_{\rm II}}(V_2)$ be DS-monotonic. From previous arguments, we know that $I\left\{\hat{P}_{x_{\rm I}^*}(V_2)\right\} < I\left\{\hat{P}_{x_{\rm I}}(V_2)\right\}$ and $I\left\{\hat{P}_{x_{\rm II}^*}(V_2)\right\} < I\left\{\hat{P}_{x_{\rm II}}(V_2)\right\}$. Then, using Lemma A.3, it can be shown that $I\left\{\hat{P}_{(x_{\rm I}^*,x_{\rm II}^*)}(V_2)\right\} < I\left\{\hat{P}_{(x_{\rm I},x_{\rm II})}(V_2)\right\}$. Thus, in this case, x^* is, indeed, $(x_{\rm I}^*, x_{\rm II}^*)$.
- Case 2: Let $\hat{P}_{x_{\rm I}}(V_2)$ and $\hat{P}_{x_{\rm II}}(V_2)$ be non-monotonic due to all active sensors of either or both states, $x_{\rm I}$ and $x_{\rm II}$, being captured. Then, it can be shown using Lemma A.5 that $I\left\{\hat{P}_{(x_{\rm I}^*,x_{\rm II}^*)}(V_2)\right\} < I\left\{\hat{P}_{(x_{\rm I},x_{\rm II})}(V_2)\right\}$. Thus, in this case also, x^* is, indeed, $(x_{\rm I}^*, x_{\rm II}^*)$.

From these two cases, it is clear that $x^* = (x_{\rm I}^*, x_{\rm II}^*)$. Therefore, $\hat{P}_{(x_{\rm I}^*,x_{\rm II}^*)}(V_2)$ is reported as the pmf of V_2.

Similarly, it can be shown that the pmf's of V_1, calculated using DPVA-I and CPVA, are the same. □

REFERENCES

[1] H. E. Garcia, N. Jhamaria, H. Kuang, W.-C. Lin, and S. M. Meerkov, "Resilient Monitoring System: Design and Performance Analysis," in *Proc. 4th Int. Symp. on Resilient Control Systems*, Boise, Idaho, USA, Aug. 9-11, 2011, pp. 61–68.

[2] H. E. Garcia, W.-C. Lin, S. M. Meerkov, and M. T. Ravichandran, "Data Quality Assessment: Modelling and Application in Resilient Monitoring Systems," in *Proc. 5th Int. Symp. on Resilient Control Systems*, Salt Lake City, Utah, USA, Aug. 14-16, 2012, pp. 124–129.

[3] ——, "Resilient monitoring systems: Architecture, design, and application to boiler/turbine plant," *submitted to IEEE Trans. on Cybernetics*, April, 2013.

[4] S. M. Meerkov, "Mathematical theory of behavior," *Mathematical Biosciences*, vol. 43, 1979.

[5] P. T. Kabamba, W.-C. Lin, and S. M. Meerkov, "Rational probabilistic deciders - Part I: Individual behavior," *Mathematical Problems in Engineering*, no. 35897, pp. 1–31, 2007.

[6] ——, "Rational probabilistic deciders - Part II: Collective behavior," *Mathematical Problems in Engineering*, no. 82184, pp. 1–34, 2007.

[7] D. D. Siljak and M. B. Vukcevic, "Decentralization, stabilization, and estimation in large-scale systems," *IEEE Trans. on Automatic Control*, vol. AC-21, pp. 363–366, 1976.

[8] D. D. Siljak, *Decentralized Control of Complex Systems*. Dover Publications, NY, 2012.

[9] G. Shafer, *A Mathematical Theory of Evidence*. Princeton University Press, 1976.

[10] S. Kullback, *Information Theory and Statistics*. John Wiley and Sons, NY, 1959.

[11] Y. Peng, S. Zhang, and R. Pan, "Bayesian network reasoning with uncertain evidences," *International Journal of Uncertainty, Fuzziness, and Knowledge-based Systems*, vol. 18, no. 5, pp. 539–564, 2010.