DISTRIBUTED ESTIMATION UNDER NETWORK MODEL UNCERTAINTY

Saurabh Sihag and Ali Tajer

Electrical, Computer, and Systems Engineering Department
Rensselaer Polytechnic Institute

ABSTRACT

This paper considers the problem of distributed state estimation in an interconnected network, in which there is uncertainty in the true model. Such uncertainties are due to the possibility of disruptions or changes in the nominal model. The focus is on the setting in which the true network model belongs to a set of possible models. Forming an optimal estimate has high computational complexity in large networks and, therefore, this paper treats this problem in a distributed framework. The key observation is that the estimation quality critically depends on successful isolation of the true model. On the other hand, the true model cannot be determined perfectly due to noisy measurements. Based on these observations, this paper formulates a composite hypotheses testing problem and provides optimal decision rules that account for estimation quality and detection performance. The theory developed in this paper is evaluated via a case study.

Index Terms— State estimation, distributed estimation, model uncertainty, networks

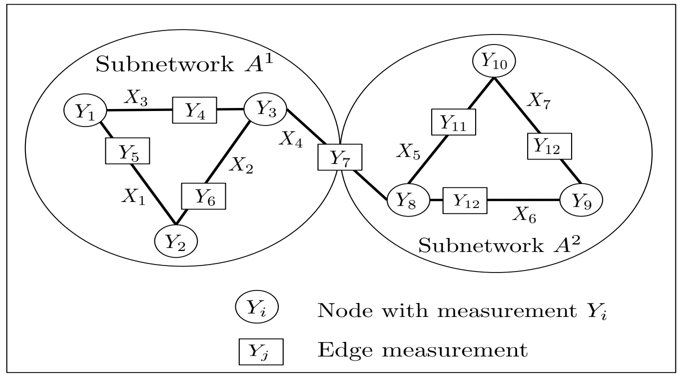

Fig. 1. Network model with X_4 as shared state

1. INTRODUCTION

1.1. Overview

Consider the problem of state estimation in an interconnected network of n agents with m connections between the agents. The agents are grouped into K interconnected subnetworks that are connected through tie-lines. Figure 1 depicts a network of $n = 6$ nodes with $m = 7$ connections that is divided into $K = 2$ subnetworks. We adopt a graph to represent the network, in which the nodes and edges represent the agents and their interconnections, respectively. We also define A^k as the set of nodes that belong to subnetwork k, i.e.,

$$A^k \triangleq \{i : \text{ node } i \text{ belongs to subnetwork } k\} \ . \tag{1}$$

To each edge $j \in \{1, \dots, m\}$ we assign a state parameter denoted by $X_j \in \mathbb{R}$, which can model, for instance, the flow between the nodes it is connecting. Accordingly, we define $\boldsymbol{X} \triangleq [X_1, \dots, X_m]$ as the state vector. The estimation objective is to form an estimate for $\boldsymbol{X}$ based on the measurements made by $(m + n)$ measurement units placed at the nodes and edges. We denote the measurements collected across the network by $\boldsymbol{Y} \in \mathbb{R}^{m+n}$, which are related to the state parameters via

$$\boldsymbol{Y} = h_0(\boldsymbol{X}) + \boldsymbol{N} \ , \tag{2}$$

where h_0 captures the networks dynamics and $\boldsymbol{N}$ accounts for the measurement noise. Corresponding to each subnetwork $k \in \{1, \dots, K\}$, we define $\boldsymbol{X}^k$ as the vector of state parameters associated with the edges that belong to this subnetwork. Specifically, $\boldsymbol{X}^k = \boldsymbol{X} \circ \mathbb{1}\{A^k\}$, where $\circ$ denotes Hadamard product, and corresponding to $A \subseteq \{1, \dots, m\}$, vector $\mathbb{1}\{A\} \in \mathbb{R}^m$ is defined such that

$$[\mathbb{1}\{A\}]_i = \begin{cases} 1 & \text{if} \quad i \in A \\ 0 & \text{if} \quad i \notin A \end{cases} \ , \tag{3}$$

where ϕ is the empty set. Clearly some state parameters belong to more than one subnetwork (e.g., X_4 in Fig. 1). Similarly, for any subnetwork $k \in \{1, \dots, K\}$, we define $\boldsymbol{Y}^k$ as the vector of measurements corresponding to the nodes and edges that belong to subnetwork A^k, i.e., $\boldsymbol{Y}^k = \boldsymbol{Y} \circ \mathbb{1}\{A^k\}$. It is noteworthy that the measurements from the tie-lines belong to more than one subnetwork (e.g., Y_7 in Fig. 1). In each subnetwork A^k, $\boldsymbol{X}^k$, and $\boldsymbol{Y}^k$ are related via

$$\boldsymbol{Y}^k = h_0^k(\boldsymbol{X}^k) + \boldsymbol{N}^k \ , \tag{4}$$

where h_0^k captures the internal dynamics of subnetwork A^k and $\boldsymbol{N}^k$ is the corresponding measurement noise. Under the condition of observability, state estimation can be performed independently by each subnetwork in the network. However, in an interconnected model, the state vectors for neighboring subnetworks overlap because of sharing of states corresponding to the tie-lines (e.g., X_4 in Fig. 1). Therefore, performing state estimation in different subnetworks independently is not necessarily optimal. Specifically, using independent state estimators for individual subnetworks does not incorporate the disparity between the estimates of the shared states. Another dimension to this state estimation problem is added when there is the possibility that network dynamics undergo stochastic changes due to, e.g., changes in network topology or malfunction in the measurement units. We assume that the nominal model might change to $s \in \mathbb{N}$ other possible models such that each subnetwork remains observable and the tie-lines are unbroken. Under the model $i \in \{1, \dots s\}$, the relationships in (2) and (4) change to

$$\boldsymbol{Y} = h_i(\boldsymbol{X}) + \boldsymbol{N} \ , \quad \text{and} \quad \boldsymbol{Y}^k = h_i^k(\boldsymbol{X}^k) + \boldsymbol{N}^k \ , \tag{5}$$

This research was supported in part by the U. S. National Science Foundation under the CAREER Award ECCS-1554482 and grant DMS-1737976.

respectively. The uncertainty in the true model introduces another decision-making process to the state estimation problem. Specifically, the estimation quality strongly hinges on the successful isolation of the true model, while the detection of true model is never perfect because of noisy measurements. This premise motivates considering the detection-dependent estimation problems in this paper. Specifically, by properly selecting cost functions for the shared parameters estimators, as well as the combined estimation and detection rules, we formalize a distributed parameter estimation framework and characterize the optimal distributed decision rules.

1.2. Relevant Studies

Distributed state estimation in dynamical models using consensus based algorithms has been studied in the domains of control theory [1–4], wireless sensor networks [5–9] and power systems [10–12]. In consensus based estimation algorithms, the nodes or the constituents of the system exchange information among themselves to converge to a state estimate for the systems. In [8], a distributed implementation of Kalman filter is studied for a sparse large-scale system being monitored by a set of sensors. Distributed state estimation under similar settings have been studied as multi-area state estimation in the power systems domain in [12–16], where the power grid is divided into several areas to lower the complexity of estimation in large grids. State estimation is performed by coordination between local estimators and a centralized estimator in [13] and [14]. A distributed approach is adopted in [15] and [16], where the areas exchange their estimates for the shared states to converge to a common estimate.

Inference under model uncertainties has been studied in [5–7, 17–21]. In [5] and [6], consensus algorithms that are robust to change in network topology are developed. Under various uncertainties in the dynamic model, robust Kalman filter based algorithms are developed in [7] and [17]. In the context of power systems, in [18–23], the uncertainties are modeled as line outages. Line outage detection and topology identification problems are considered in [20], where hidden Markov models are used to model the measurements and outage detection.

In this paper, we consider a state estimation problem in the setting where the true model belongs to a fixed set of possible models. A similar setting is studied in [18] and [19], where closed-form expressions for the joint posterior probability of the true model and the states is developed and utilized for designing optimal outage detectors. Using decoupled strategies for model isolation and state estimation do not ensure optimality, as established in [24] and [25]. For example, using Neyman-Pearson detection test to identify the true model followed by optimal estimation does not integrate the imperfection in the detection step into the estimator design. In [26] and [27], optimal joint detection and estimation rules are developed under constraints on the detection power. Specifically, in [27] decision rules are developed for a binary hypotheses testing problem with unknown parameters to be estimated under both hypotheses. In this paper, we consider the problems of distributed state estimation and true model isolation jointly and model it as an $(s+1)$-composite hypotheses testing problem and provide optimal decision rules under distributed estimation framework. The decision rules are evaluated on a network in the case study.

2. ESTIMATION FRAMEWORK

Based on the definitions in Section 1.1, we assume that multiple states can be shared between any two subnetworks and, a state can be shared among more than two subnetworks. We assume that noise $\boldsymbol{N}$ is distributed according to a distribution with a known probability density function (pdf) $f_{\boldsymbol{N}}$. The state vector $\boldsymbol{X}$ has a prior known pdf π. The pdf of the measurement vector $\boldsymbol{Y}$ under model $i \in \{0, \dots, s\}$ is denoted by f_i. Also, we define ϵ_i as the prior probability of h_i representing the true model for $i \in \{0, \dots s\}$. The structure of the optimal estimator under different models is different. Let $\hat{\boldsymbol{X}}_i(\boldsymbol{Y})$ denote the estimate of $\boldsymbol{X}$ under the model h_i. The closeness of $\hat{\boldsymbol{X}}_i(\boldsymbol{Y})$ and $\boldsymbol{X}$ can be quantified by the quadratic cost function

$$\mathsf{C}(\boldsymbol{X}, \boldsymbol{U}) \triangleq \|\boldsymbol{X} - \boldsymbol{U}\|^2 \ , \tag{6}$$

for any generic estimator $\boldsymbol{U}$. The average posterior cost function for an estimate of $\boldsymbol{X}$ under model h_i, is

$$\mathsf{C}_{\rm p,i}(\boldsymbol{U} \mid \boldsymbol{Y}) \triangleq \mathbb{E}_i \left[\mathsf{C}(\boldsymbol{X}, \boldsymbol{U}) \mid \boldsymbol{Y} \right] \ , \ \text{for } i \in \{0, \dots, s\} \ . \tag{7}$$

The optimal estimator that minimizes $\mathsf{C}_{\rm p,i}(\boldsymbol{U} \mid \boldsymbol{Y})$ is the standard Bayesian minimum mean squared error (MMSE) estimator given by

$$\hat{\boldsymbol{X}}_i(\boldsymbol{Y}) \triangleq \mathbb{E}_i(\boldsymbol{X} \mid \boldsymbol{Y}) \ , \tag{8}$$

and the associated posterior estimation cost is

$$\hat{\mathsf{C}}_{\rm p,i}(\boldsymbol{Y}) \triangleq \min_{\boldsymbol{U}} \mathsf{C}_{\rm p,i}(\boldsymbol{U} \mid \boldsymbol{Y}) \ . \tag{9}$$

Since all the subnetworks are observable, the definitions in (7)-(9) can be accordingly defined for each subnetwork $k \in \{1, \dots, K\}$, and under each model h_i. Specifically, for subnetwork k, we define the average posterior cost function under model h_i as

$$\mathsf{C}^k_{\rm p,i}(\boldsymbol{U} \mid \boldsymbol{Y}^k) \triangleq \mathbb{E}_i[\mathsf{C}(\boldsymbol{X}^k, \boldsymbol{U}) \mid \boldsymbol{Y}^k] \ , \ \text{for } i \in \{0, \dots, s\} \ , \tag{10}$$

for any generic estimator $\boldsymbol{U} \in \mathbb{R}^m \circ \mathbb{1}\{A^k\}$ of $\boldsymbol{X}^k$. The optimal posterior estimation cost for subnetwork k is given by

$$\hat{\mathsf{C}}^k_{\rm p,i}(\boldsymbol{Y}^k) \triangleq \operatorname*{argmin}_{\boldsymbol{U}} \mathsf{C}^k_{\rm p,i}(\boldsymbol{U} \mid \boldsymbol{Y}^k) \ , \tag{11}$$

which is achieved by the local estimator based on the local data $\boldsymbol{Y}^k$

$$\hat{\boldsymbol{X}}^k_i(\boldsymbol{Y}^k) \triangleq \mathbb{E}_i[\boldsymbol{X}^k \mid \boldsymbol{Y}^k] \ . \tag{12}$$

Since forming an optimal centralized estimator can be computationally complex or unfeasible, we consider a distributed estimation structure that combines the local estimates from each subnetwork. Note that forming locally-optimal estimates independently of other subnetworks can be sub-optimal because it does not incorporate the dependencies among various subnetworks that have common states. To circumvent this, and in order to formalize a distributed estimation structure that incorporates the discrepancies between the estimates formed by different subnetworks for the shared states, we also provide modified posterior cost functions that signify these discrepancies. Specifically, for a generic state estimator $\boldsymbol{U}^k_i$ for $\boldsymbol{X}^k$ under model h_i, we modify $\mathsf{C}_{\rm p,i}(\boldsymbol{U} \mid \boldsymbol{Y})$ in (7) such that the fidelity of the set of estimators $\boldsymbol{U}_i = \{\boldsymbol{U}^1_i, \dots, \boldsymbol{U}^K_i\}$ is quantified by the following posterior cost function

$$\begin{aligned}\mathsf{R}_{\rm p,i}(\boldsymbol{U}_i \mid \boldsymbol{Y}) \triangleq & \sum_{k=1}^{K} \mathsf{C}^k_{\rm p,i}(\boldsymbol{U}^k \mid \boldsymbol{Y}^k) \\ & + \frac{1}{2} \sum_{k,l \in \{1,\dots K\}} \|(\boldsymbol{U}^k - \boldsymbol{U}^l) \circ \mathbb{1}\{A^k \cap A^l\}\|^2 \ ,\end{aligned} \tag{13}$$

where the first term captures the aggregate local estimation costs, and the second term adds a cost associated with the discrepancies between the estimates of the common parameters provided by different subnetworks. Under model h_i, for $i \in \{0, \dots, s\}$, the optimal posterior cost is given by

$$\mathsf{R}^*_{\rm p,i}(\boldsymbol{Y}) \triangleq \min_{\boldsymbol{U}_i} \mathsf{R}_{\rm p,i}(\boldsymbol{U}_i \mid \boldsymbol{Y}) \ . \tag{14}$$

3. PROBLEM FORMULATION

When there is the possibility that the true model of the network deviates from its nominal model, the quality of the state estimator becomes intertwined with the correct isolation of the true model. Therefore, we formalize the estimation problem of interest, as an inherently detection-dependent estimation problem, which leads to solving the following $(s+1)$-ary composite hypothesis testing problem:

$$\mathsf{H}_i: \quad \boldsymbol{Y} \sim f_i(\boldsymbol{Y} \mid \boldsymbol{X}), \text{ with } \boldsymbol{X} \sim \pi(\boldsymbol{X}) , \tag{15}$$

where hypothesis H_0 represents the nominal model, and hypothesis H_i represents model h_i, for $i \in \{1, \dots, s\}$. Note that decoupling the problem in (15) into independent detection and estimation problems does not guarantee the optimal performance because the uncertainty in the one step is not incorporated into the design in the second step (for example, Neyman-Pearson based detection followed by optimal estimation).

3.1. Definitions

Define $\boldsymbol{\delta} \triangleq [\delta_0(\boldsymbol{Y}), \dots \delta_s(\boldsymbol{Y})]$ as the randomized detection rule for the problem in (15), where $\delta_i(\boldsymbol{Y})$ represents the probability of the decision in favor of H_i, for $i \in \{0, \dots s\}$. Let $\mathsf{D} \in \{\mathsf{H}_0, \dots \mathsf{H}_s\}$ denote the decision formed and $\mathsf{T} \in \{\mathsf{H}_0, \dots \mathsf{H}_s\}$ be the true model. Then, the likelihood of forming a decision in favor of H_j when the true model is H_i is given by

$$\mathbb{P}(\mathsf{D}=\mathsf{H}_j \,|\, \mathsf{T}=\mathsf{H}_i) = \int_{\boldsymbol{Y}} \delta_j(\boldsymbol{Y}) f_i(\boldsymbol{Y})\, d\boldsymbol{Y} \ , \text{ for } i,j \in \{0, \dots, s\} . \tag{16}$$

We define $\mathsf{P}_{\rm md}$ as the error rate when the true model is not the nominal model

$$\mathsf{P}_{\rm md} \triangleq \mathbb{P}(\mathsf{D} \neq \mathsf{T} \,|\, \mathsf{T} \neq \mathsf{H}_0) \tag{17}$$

$$= \frac{1}{\mathbb{P}(\mathsf{T} \neq \mathsf{H}_0)} \sum_{i=1}^{s} \mathbb{P}(\mathsf{D} \neq \mathsf{H}_i \,|\, \mathsf{T}=\mathsf{H}_i) \mathbb{P}(\mathsf{T}=\mathsf{H}_i) . \tag{18}$$

Using the definition of ϵ_i and (16), we can rewrite $\mathsf{P}_{\rm md}$ as

$$\mathsf{P}_{\rm md} = \sum_{i=0}^{s} \sum_{\substack{j=0 \\ j \neq i}}^{s} \frac{\epsilon_i}{1-\epsilon_0} \int_{\boldsymbol{Y}} \delta_j(\boldsymbol{Y}) f_i(\boldsymbol{Y})\, d\boldsymbol{Y} . \tag{19}$$

We also define $\mathsf{P}_{\rm fa}$ as the error rate under the nominal model, i.e.,

$$\mathsf{P}_{\rm fa} \triangleq \mathbb{P}(\mathsf{D} \neq \mathsf{T} \,|\, \mathsf{T}=\mathsf{H}_0) = \sum_{i=1}^{s} \mathbb{P}(\mathsf{D}=\mathsf{H}_i \,|\, \mathsf{T}=\mathsf{H}_0) ,$$

$$= \sum_{i=1}^{s} \int_{\boldsymbol{Y}} \delta_i(\boldsymbol{Y}) f_0(\boldsymbol{Y})\, d\boldsymbol{Y} . \tag{20}$$

For a set of generic estimators $\boldsymbol{U}_i \triangleq \{\boldsymbol{U}_i^1, \dots, \boldsymbol{U}_i^K\}$, where $\boldsymbol{U}_i^k$ is the local estimator for subnetwork k under model H_i for $i \in \{0, \dots, s\}$, the corresponding estimation cost is given by $\mathsf{C}(\boldsymbol{X}, \boldsymbol{U}_i)$. Note that the estimation cost $\mathsf{C}(\boldsymbol{X}, \boldsymbol{U}_i)$ is relevant only when the decision is in the favor of H_i. Therefore, we define $J_i(\delta_i, \boldsymbol{U}_i)$ as the average estimation cost given that the decision is H_i under hypothesis H_i, i.e.,

$$J_i(\delta_i, \boldsymbol{U}_i) \triangleq \frac{\int_{\boldsymbol{Y}} \delta_i(\boldsymbol{Y}) \mathsf{R}_{\rm p,i}(\boldsymbol{U}_i \mid \boldsymbol{Y}) f_i(\boldsymbol{Y})\, d\boldsymbol{Y}}{\int_{\boldsymbol{Y}} \delta_i(\boldsymbol{Y}) f_i(\boldsymbol{Y})\, d\boldsymbol{Y}} . \tag{21}$$

The overall estimation cost is defined as the maximum of the average costs $J_i(\delta_i, \boldsymbol{U}_i)$, i.e.,

$$J(\boldsymbol{\delta}, \boldsymbol{U}) \triangleq \max_{i \in \{0, \dots s\}} J_i(\delta_i, \boldsymbol{U}_i) , \tag{22}$$

where $\boldsymbol{U} \triangleq [\boldsymbol{U}_1, \dots, \boldsymbol{U}_s]$.

3.2. Optimal Detection-dependent Estimation

In this section, we design the estimation framework for optimizing the state estimation quality under uncertainty in the true model. Because of the presence of noise in the measurements, the detection is not perfect. At the same time, the estimation quality depends strongly on the detection accuracy of the true model. Therefore, we aim to jointly design the estimators and the detection rules under the constraints on error rates $\mathsf{P}_{\rm fa}$ and $\mathsf{P}_{\rm md}$, i.e.,

$$\mathcal{P}(\alpha, \beta) \triangleq \begin{cases} \min_{(\boldsymbol{\delta}, \boldsymbol{U})} & J(\boldsymbol{\delta}, \boldsymbol{U}) \\ \text{s.t.} & \mathsf{P}_{\sf md} \leq \beta \\ & \mathsf{P}_{\sf fa} \leq \alpha \end{cases} , \tag{23}$$

where $\alpha, \beta \in (0, 1)$.

Remark 1 (Feasibility)**.** *The constraints on $\mathsf{P}_{\rm fa}$ and $\mathsf{P}_{\rm md}$ cannot be made arbitrarily small simultaneously because according to Neyman-Pearson lemma, there exists a minimum feasible value $\beta^*(\alpha)$ for $\mathsf{P}_{\rm md}$ for a given constraint α on $\mathsf{P}_{\rm fa}$. Therefore, β must be chosen such that $\beta \geq \beta^*(\alpha)$.*

4. STATE ESTIMATORS UNDER MODEL UNCERTAINTY

By using the expansions of $\mathsf{P}_{\rm md}$ and $\mathsf{P}_{\rm fa}$ in (19) and (20), we can write the problem in (23) as

$$\mathcal{P}(\alpha, \beta) = \begin{cases} \min_{(\boldsymbol{\delta}, \boldsymbol{U})} & J(\boldsymbol{\delta}, \boldsymbol{U}) \\ \text{s.t.} & \sum\limits_{i=1}^{s} \sum\limits_{\substack{j=0, \\ j \neq i}}^{s} \frac{\epsilon_i}{1-\epsilon_0} \int_{\boldsymbol{Y}} \delta_j(\boldsymbol{Y}) f_i(\boldsymbol{Y}) d\boldsymbol{Y} \leq \beta \\ & \sum\limits_{i=1}^{s} \int_{\boldsymbol{Y}} \delta_i(\boldsymbol{Y}) f_0(\boldsymbol{Y})\, d\boldsymbol{Y} \leq \alpha \end{cases} . \tag{24}$$

Note that the estimators appear only in the estimation cost $J(\boldsymbol{\delta}, \boldsymbol{U})$, which allows for decoupling the problem in (24) into two subproblems, as stated in the following theorem.

Theorem 1. *The problem $\mathcal{P}(\alpha, \beta)$ can be equivalently written in the following form*

$$\mathcal{P}(\alpha, \beta) = \begin{cases} \min_{\boldsymbol{\delta}} & \tilde{J}(\boldsymbol{\delta}, \hat{\boldsymbol{X}}) \\ \text{s.t.} & \sum\limits_{j=1}^{s} \sum\limits_{\substack{i=0, \\ i \neq j}}^{s} \frac{\epsilon_j}{1-\epsilon_0} \int_{\boldsymbol{Y}} \delta_i(\boldsymbol{Y}) f_j(\boldsymbol{Y}) d\boldsymbol{Y} \leq \beta \\ & \sum\limits_{i=1}^{s} \int_{\boldsymbol{Y}} \delta_i(\boldsymbol{Y}) f_0(\boldsymbol{Y})\, d\boldsymbol{Y} \leq \alpha \end{cases} , \tag{25}$$

$$\textit{where} \qquad \hat{\boldsymbol{X}} \triangleq \arg\min_{\boldsymbol{U}} J(\boldsymbol{\delta}, \boldsymbol{U}) , \tag{26}$$

$$\textit{and} \qquad \tilde{J}(\boldsymbol{\delta}, \hat{\boldsymbol{X}}) \triangleq \min_{\boldsymbol{U}} J(\boldsymbol{\delta}, \boldsymbol{U}) . \tag{27}$$

The design of the optimal estimators is provided in Theorem 2.

Theorem 2 (State Estimator)**.** *The estimator that solves the problem in* (26) *is given by*

$$\hat{\boldsymbol{X}}_i(\boldsymbol{Y}) = \arg\inf_{\boldsymbol{U}_i} \mathsf{R}_{\rm p,i}(\boldsymbol{U}_i \mid \boldsymbol{Y}) . \tag{28}$$

As a result, the cost function $\tilde{J}(\boldsymbol{\delta}, \hat{\boldsymbol{X}})$ is given by

$$\tilde{J}(\boldsymbol{\delta}, \hat{\boldsymbol{X}}) = \max_{i \in \{0, \dots, s\}} \left\{ \frac{\displaystyle\int_{\boldsymbol{Y}} \delta_i(\boldsymbol{Y}) \mathsf{R}^*_{\rm p,i}(\boldsymbol{Y}) f_i(\boldsymbol{Y})\, d\boldsymbol{Y}}{\displaystyle\int_{\boldsymbol{Y}} \delta_i(\boldsymbol{Y}) f_i(\boldsymbol{Y})\, d\boldsymbol{Y}} \right\} . \tag{29}$$

In the following theorem, we provide closed-form expressions for $\hat{\boldsymbol{X}}_i(\boldsymbol{Y})$. Specifically, the structure of the optimal state estimator that minimizes the cost function defined in (13) is provided in Theorem 3. To formalize the results, define $\mathcal{S}_l$ as the set of subnetworks that X_l belongs to, i.e.,

$$\mathcal{S}_l \triangleq \{k : l \in A^k\} \ , \quad \text{for} \quad l \in \{1, \dots m\} \ . \tag{30}$$

Theorem 3. *Denote the optimal estimate of X_l in subnetwork $A^k, k \in \mathcal{S}_l$ by $\hat{X}_l^k$. Under model h_j, for $j \in \{0, \dots, s\}$, the optimal state estimates for the state parameter X_l in subnetworks $\mathcal{S}_l$, for $l \in \{1, \dots m\}$, satisfy*

$$\forall k \in \mathcal{S}_l : \hat{X}_l^k \,(1 + |\mathcal{S}_l|) = \mathbb{E}_j[X_l \,|\, \boldsymbol{Y}^k] + \sum_{r \in \mathcal{S}_l} \hat{X}_l^r \ . \tag{31}$$

This theorem indicates that the local estimates of X_l by the subnetwork that share it can be found by solving an $|\mathcal{S}_l|$-dimensional linear system with the $|\mathcal{S}_l|$ equalities specified by (31). These local estimates collectively minimize the cost function defined in (13). Using the optimal estimators designed in Theorem 2, the corresponding optimal detection rules are designed by solving (25).

Theorem 4 (Detection and Isolation)**.** *The optimal decision rule is $\delta_{i^*}(\boldsymbol{Y}) = 1$, where $i^* \triangleq \operatorname{argmin}_{i \in \{0, \dots, s\}} B_i$. B_0 is defined as*

$$B_0 \triangleq \ell_0 f_0(\boldsymbol{Y})(\mathsf{R}^*_{\mathrm{p},0}(\boldsymbol{Y}) - u) + \ell_{s+1} \sum_{i=1}^{s} \frac{\epsilon_i}{1-\epsilon_0} f_i(\boldsymbol{Y}) \ , \tag{32}$$

and $\{B_i : i \in \{1, \dots, s\}\}$ are defined as

$$B_i \triangleq \ell_i f_i(\boldsymbol{Y})(\mathsf{R}^*_{\mathrm{p,i}}(\boldsymbol{Y}) - u) + \ell_{s+1} \sum_{\substack{j=1,\\ j \neq i}}^{s} \frac{\epsilon_j}{1-\epsilon_0} f_j(\boldsymbol{Y}) + \ell_{s+2} f_0(\boldsymbol{Y}) \ , \tag{33}$$

and the non-negative constants $\{\ell_i : i \in \{0, \dots, s+2\}\}$ are the Lagrangian multipliers selected such that $\sum_{i=0}^{s+2} \ell_i = 1$, and the constraints in the following convex optimization problem (which is equivalent to the problem in (25)) are satisfied.

$$\tilde{\mathcal{P}}(\alpha, \beta) = \begin{cases} \min_{\delta} & u \\ \text{s.t.} & \displaystyle\int_{\boldsymbol{Y}} \delta_i(\boldsymbol{Y}) f_i(\boldsymbol{Y})(\mathsf{R}^*_{\mathrm{p,i}}(\boldsymbol{Y}) - u) d\boldsymbol{Y} \le 0 \\ & \displaystyle\sum_{j=1}^{s} \sum_{\substack{i=0 \\ i \neq j}}^{s} \frac{\epsilon_j}{1-\epsilon_0} \int_{\boldsymbol{Y}} \delta_i(\boldsymbol{Y}) f_j(\boldsymbol{Y}) \, d\boldsymbol{Y} \ \le \beta \\ & \displaystyle\sum_{i=1}^{s} \int_{\boldsymbol{Y}} \delta_i(\boldsymbol{Y}) f_0(\boldsymbol{Y}) \, d\boldsymbol{Y} \le \alpha \end{cases} . \tag{34}$$

5. CASE STUDY

We evaluate the estimation quality obtained by the design of the state estimators and decision rules established in Section 4 on the network model in Fig. 1. We assume that the states X_i, for $i \in \{1, \dots, 7\}$ are independently and identically distributed (i.i.d) with pdf $\text{Unif}[0,2]$, and the noise component in all measurements is i.i.d with pdf $\mathcal{N}(0, 0.1)$. Of all the possible measurements available in a subnetwork, we assume that the subnetwork can choose a subset of them such that it is observable. In subnetwork A^1, we assume that the measurements $\boldsymbol{Y}^1 = [Y_1, Y_2, Y_3, Y_4, Y_5]$ are related to the states $\boldsymbol{X}^1 = [X_1, X_2, X_3, X_4]$ as

$$\boldsymbol{Y}^1 = \boldsymbol{H}^1 \boldsymbol{X}^1 + \boldsymbol{N}^1 \ , \tag{35}$$

where

$$\boldsymbol{H}^1 = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} .$$

Similarly, in subnetwork A^2, the measurements $\boldsymbol{Y}^2 = [Y_8, Y_9, Y_{10}, Y_{11}, Y_{12}]$ are related to the states $\boldsymbol{X}^2 = [X_4, X_5, X_6, X_7]$ according to

$$\boldsymbol{Y}^2 = \boldsymbol{H}^2 \boldsymbol{X}^2 + \boldsymbol{N}^2 \ , \tag{36}$$

where

$$\boldsymbol{H}^2 = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} .$$

We assume that the network model can be altered by two mutually exclusive events:

1. Malfunction in the node with measurement Y_1, with probability $\epsilon_1 = 0.3$, such that when malfunctioning, Y_1 is only noise, and
2. Malfunction in the node with measurement Y_{10}, with probability $\epsilon_2 = 0.2$, such that when malfunctioning, Y_{10} is only noise.

We denote the ratio of $\mathcal{P}(\alpha, \beta)$ and the average estimation cost under no uncertainty by q and evaluate it for different values of α and β.

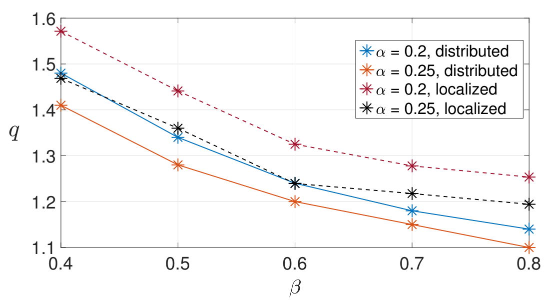

Fig. 2. q versus β for different α

Figure 2 shows the characterization of the cost $\mathcal{P}(\alpha, \beta)$ normalized by the estimation cost when there is no model uncertainty and its comparison with the case when the state estimates of shared states are calculated independently by the subnetworks locally. As the constraints on α and β are relaxed, the estimation cost decreases monotonically. Also, the distributed estimation strategy outperforms the localized estimation strategy.

6. CONCLUSION

In this paper, we have considered the problem of distributed state estimation in an interconnected network where we face uncertainty in the network model. The uncertainty in the true model is formulated as the possibility of the true model deviating from the nominal model to a set of alternative models with prior probabilities. Under this setting, we have formulated the distributed state estimation problem jointly with the design for detection rules for isolating the true model. We have shown that the detection and estimation routines are intertwined and have provided the design for optimal estimators and detection rules. We have also evaluated the theory developed in a case study and observed that the estimation strategy proposed in this paper performs better than the estimation strategy based where the shared states are estimated independently by the subnetworks.

7. REFERENCES

[1] J. Shamma, *Cooperative control of distributed multi-agent systems*. John Wiley and Sons, 2008.

[2] R. Olfati-Saber, J. A. Fax, and R. M. Murray, “Consensus and cooperation in networked multi-agent systems,” *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, Jan. 2007.

[3] A. Bemporad, M. Heemels, and M. Johansson, *Networked control systems*. Springer, 2010, vol. 406.

[4] W. Ren, R. W. Beard, and D. B. Kingston, “Multi-agent Kalman consensus with relative uncertainty,” in *Proc. American Control Conference*, Portland, OR, Jun. 2005, pp. 1865–1870.

[5] L. Xiao, S. Boyd, and S. Lall, “A scheme for robust distributed sensor fusion based on average consensus,” in *Proc. International symposium on Information processing in sensor networks*, Los Angeles, CA, Apr. 2005, p. 9.

[6] S. Kar and J. M. F. Moura, “Distributed consensus algorithms in sensor networks with imperfect communication: Link failures and channel noise,” *IEEE Transactions on Signal Processing*, vol. 57, no. 1, pp. 355–369, Jan. 2009.

[7] B. Chen, W. Zhang, G. Hu, and L. Yu, “Networked fusion Kalman filtering with multiple uncertainties,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51, no. 3, pp. 2232–2249, Jul. 2015.

[8] U. A. Khan and J. M. Moura, “Distributing the Kalman filter for large-scale systems,” *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 4919–4935, Oct. 2008.

[9] S. Kar, J. M. Moura, and K. Ramanan, “Distributed parameter estimation in sensor networks: Nonlinear observation models and imperfect communication,” *IEEE Transactions on Information Theory*, vol. 58, no. 6, pp. 3575–3605, Jun. 2012.

[10] X. Qing, H. R. Karimi, Y. Niu, and X. Wang, “Decentralized unscented Kalman filter based on a consensus algorithm for multi-area dynamic state estimation in power systems.” *International Journal of Electrical Power and Energy Systems*, vol. 65, pp. 26–33, Feb. 2015.

[11] L. Xie, D.-H. Choi, S. Kar, and H. V. Poor, “Fully distributed state estimation for wide-area monitoring systems,” *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1154–1169, Sep. 2012.

[12] X. Li and A. Scaglione, “Robust decentralized state estimation and tracking for power systems via network gossiping.” *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 7, pp. 1184–1194, Jul. 2013.

[13] L. Zhao and A. Abur, “Multi area state estimation using synchronized phasor measurements.” *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 611–617, May 2005.

[14] G. N. Korres, “A distributed multiarea state estimation.” *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 73–84, Apr. 2011.

[15] V. Kekatos and G. B. Giannakis., “Distributed robust power system state estimation,” *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1617–1626, Oct. 2013.

[16] A. J. Conejo, S. de la Torre, and M. Canas, “An optimization approach to multiarea state estimation,” *IEEE Transactions on Power Systems*, vol. 22, no. 1, pp. 213–221, Feb. 2007.

[17] P. Ren, H. Lev-Ari, and A. Abur, “Robust continuous-discrete extended Kalman filter for estimating machine states with model uncertainties,” in *Proc. Power Systems Computation Conference*, Genoa, Italy, 2016, pp. 1–7.

[18] Y. Zhao, J. Chen, A. Goldsmith, and H. V. Poor, “Identification of outages in power systems with uncertain states and optimal sensor locations.” *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 6, pp. 1140–1153, Nov. 2014.

[19] ——, “Dynamic joint outage identification and state estimation in power systems,” in *Proc. Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, Nov. 2014, pp. 1138–1142.

[20] Q. Huang, L. Shao, and N. Li, “Dynamic detection of transmission line outages using hidden Markov models,” *IEEE Transactions on Power Systems*, vol. 31, no. 3, pp. 2026–2033, Aug. 2016.

[21] Y. Zhao, A. Goldsmith, and H. V. Poor., “On PMU location selection for line outage detection in wide-area transmission networks,” in *Proc. Power and Energy Society General Meeting*, San Diego, CA, Jul. 2012, pp. 1–8.

[22] A. Tajer, S. Kar, H. V. Poor, and S. Cui, “Distributed joint cyber attack detection and state recovery in smart grids,” in *Proc. IEEE International Conference on Smart Grid Communications*, Brussels, Belgium, Oct. 2011, pp. 202–207.

[23] A. Tajer, “State recovery under cyber attacks in energy grids,” in *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop*, La Coruna, Spain, Jun. 2014, pp. 65–68.

[24] D. Middleton and R. Esposito, “Simultaneous optimum detection and estimation of signals in noise,” *IEEE Transactions on Information Theory*, vol. 14, no. 3, pp. 434–444, May 1968.

[25] O. Zeitouni, J. Ziv, and N. Merhav, “When is the generalized likelihood ratio test optimal?” *IEEE Transactions on Information Theory*, vol. 38, no. 5, pp. 1597–1602, Sep. 1992.

[26] G. V. Moustakides, G. H. Jajamovich, A. Tajer, and X. Wang, “Joint detection and estimation: Optimum tests and applications,” *IEEE Transactions on Information Theory*, vol. 58, no. 7, pp. 4215–4229, Jul. 2012.

[27] G. H. Jajamovich, A. Tajer, and X. Wang, “Minimax-optimal hypothesis testing with estimation-dependent costs,” *IEEE Transactions on Signal Processing*, vol. 60, no. 12, pp. 6151–6165, Dec. 2012.