Optimal Sensor Placement for Kalman Filtering in Stochastically Forced Consensus Networks

Lintao Ye, Sandip Roy and Shreyas Sundaram

***Abstract*— Given a linear dynamical system affected by noise, we consider the problem of optimally placing sensors (at design-time) subject to certain budget constraints to minimize the trace of the steady-state error covariance of the Kalman filter. Previous work has shown that this problem is NP-hard in general. In this paper, we impose additional structure by considering systems whose dynamics are given by a stochastic matrix corresponding to an underlying consensus network. In the case when there is a single input at one of the nodes in a tree network, we provide an optimal strategy (computed in polynomial-time) to place the sensors over the network. However, we show that when the network has multiple inputs, the optimal sensor placement problem becomes NP-hard.**

I. Introduction

In large-scale control system design, one of the key problems is to place the sensors or actuators of the system in order to achieve certain performance criteria (e.g., [1], [2]). In cases involving linear systems with process or measurement noise, researchers have studied how to place sensors in order to minimize certain metrics of the error covariance of the corresponding Kalman filter (e.g., [3], [4], [5], [6], [7]). The problem was shown to be NP-hard and inapproximable within any constant factor [8]. Thus, in this paper, we turn to systems with special properties to seek polynomial-time algorithms for the optimal sensor placement problem.

More specifically, we consider a discrete-time linear dynamical system whose states represent nodes in an undirected consensus network, and interact according to the topology of the network. Each node of the network is possibly affected by a Gaussian input. This kind of consensus system with stochastic inputs has received much attention from researchers recently (e.g., [9], [10], [11], [12]). Given a consensus system with stochastic inputs, we seek a graph-theoretic approach to optimally place sensors at the nodes of the associated graph to optimize the steady-state error covariance matrix of the corresponding Kalman filter. We refer to this problem as the *Graph-based Kalman Filtering Sensor Placement (GKFSP)* problem. We summarize some related work as follows.

In papers [6] and [8], the authors considered the same general problem as we consider here, and showed that this problem is NP-hard and that there is no constant-factor polynomial-time approximation algorithm for this problem. However, neither of them considered systems evolving over a network. In contrast, we impose additional structure by considering the system dynamics matrix to be a stochastic matrix defined over a network with Gaussian inputs and seek polynomial-time algorithms for this problem.

In [5] and [7], the authors considered the optimal sensor placement problem for Kalman filtering over a finite time interval and provided near-optimal greedy algorithms to optimize certain metrics of the error covariance of the Kalman filter. However, their results cannot be directly applied to the problem we consider here, since we aim to optimize the steady-state error covariance of the Kalman filter.

In papers [2], [13] and [14], the authors considered systems evolving over networks. In [2], the authors considered generic structural systems [15], and in [13] and [14], the authors considered consensus systems with (unknown) inputs and focused on analyzing how the inputs affect the consensus computation of the network. In this paper, we consider consensus systems with specific parameters, and consider the presence of Gaussian inputs, which then motivates the need for optimal sensor placement for Kalman filtering.

The authors in [9] and [10] considered the leader selection problem in consensus systems with stochastic inputs. The problem is to select an optimal subset of nodes whose states are fixed over time such that the H_2 norm of the system states at steady state is minimized. In contrast, we study the problem of selecting an optimal subset of nodes at which to place sensors such that the steady-state error covariance of the Kalman filter is optimized.

Our contributions are summarized as follows. First, we consider the case when there is a single stochastic input in the consensus system corresponding to a tree, and obtain the steady-state error covariance of the corresponding Kalman filter by establishing a relationship between the Kalman filter and a (delayed) unknown input observer of the system. We then provide an optimal polynomial-time algorithm for the GKFSP problem. Our second contribution is to show that when the consensus system has multiple inputs, the corresponding GKFSP problem becomes NP-hard.

A. Notation and terminology

The set of natural numbers, integers, real numbers, rational numbers, and complex numbers are denoted as $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{Q}$, and $\mathbb{C}$, resp. For a matrix $P \in \mathbb{R}^{n\times n}$, denote P^T as its transpose, P_{ij} as the element in the ith row and jth column of P, $\rho(P)$ as its spectral radius, and $P_{(i,j)}$ as a submatrix of P obtained from P by removing the ith row and jth column of P. The set of n by n positive-semidefinite matrices is denoted by $\mathbb{S}^n_+$. The identity matrix

This research was supported by NSF grant CMMI-1635014. Lintao Ye and Shreyas Sundaram are with the School of Electrical and Computer Engineering at Purdue University. Email: {ye159,sundara2}@purdue.edu. Sandip Roy is with the School of Electrical Engineering and Computer Science at Washington State University. Email: sroy@eecs.wsu.edu.

of dimension n is denoted as I_n. For a vector f, denote its ith element as f_i, and let supp(f) be its support, where supp$(f) = \{i : f_i \neq 0\}$. Denote the Euclidean norm of f by $\|f\|$. Define $\mathbf{e}_i$ to be a column vector where the ith element is 1 and all the other elements are zero; the dimension of the vector can be inferred from the context. Denote $\mathbb{E}[x]$ as the expectation of a random variable (vector) x. For a set $\mathcal{A}$, let $|\mathcal{A}|$ be its cardinality. Denote $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ as an (undirected) graph with vertex set $\mathcal{V}$ and edge set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. Each edge $(i, j) \in \mathcal{E}$ has an associated weight $m_{ij} \in \mathbb{R}_{>0}$; take $m_{ij} = 0$ if $(i, j) \notin \mathcal{E}$ and assume that $m_{ij} = m_{ji}$ for all $i \neq j$. The neighbors of vertex $i \in \mathcal{V}$ in graph $\mathcal{G}$ are given by the set $\mathcal{N}_i \triangleq \{j \in \mathcal{V} \big| (i, j) \in \mathcal{E}\}$. The degree of vertex i is defined as $d_i = \sum_{j \in \mathcal{N}_i} m_{ij}$, and the maximum degree of the vertices in the graph is defined as $d_{\max}$. Denote p_{ij} as the shortest path from vertex i to j in graph $\mathcal{G}$, and denote l_{ij} as its length (also known as the distance between i and j), obtained after all the edge weights are set to be one.

II. Problem Formulation

Consider a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with $\mathcal{V} = \{1, \ldots, n\}$. The (weighted) adjacency matrix for the graph is a matrix $M \in \mathbb{R}^{n \times n}$, where $M_{ij} = m_{ij}$ for all $i, j \in \mathcal{V}$. The degree matrix is $D \triangleq \mathrm{diag}(d_1, \ldots, d_n)$. The Laplacian matrix for the graph is then given by $L(\mathcal{G}) = D - M$.

Definition 1: Given a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with the symmetric adjacency matrix $M \in \mathbb{R}^{n \times n}$, the matrix A is defined as $A = I_n - \epsilon L(\mathcal{G})$, where $\epsilon \in (0, 1/d_{\max}]$ and $L(\mathcal{G})$ is the Laplacian matrix for the graph $\mathcal{G}$. □

Remark 1: Note that A is a symmetric and (doubly) stochastic matrix. For any $v_i, v_j \in \mathcal{V}$ such that $i \neq j$, $A_{ij} > 0$ if and only if $(v_i, v_j) \in \mathcal{E}$; for all $i \in \mathcal{V}$, $A_{ii} \geq 0$. □

Consider the matrix A as described in Definition 1 and the set of vertices that have inputs, denoted as $\mathcal{I} \triangleq \{i_0, \ldots, i_{\tau-1}\} \subseteq \mathcal{V}$, where $\tau \in \mathbb{Z}_{\geq 1}$. We consider the discrete-time linear system

$$x_{k+1} = Ax_k + Bw_k, \tag{1}$$

where $x_k \in \mathbb{R}^n$ is the system state and $(x_k)_i$ is associated with vertex $i \in \mathcal{V}$, and $B \triangleq \begin{bmatrix} B_0 & \cdots & B_{\tau-1} \end{bmatrix}$ is the input matrix, where $B_j = \mathbf{e}_{i_j}$ if there is an input at vertex $i_j \in \mathcal{I}$ for $j = 0, \ldots, \tau - 1$. The system noise $w_k \in \mathbb{R}^\tau$ is a zero-mean white Gaussian noise process with $\mathbb{E}[w_k(w_k)^T] = W$ for all $k \in \mathbb{N}$. The initial state x_0 is assumed to be Gaussian with mean $\bar{x}_0$ and covariance Π_0, and is also assumed to be independent of w_k for all $k \geq 0$.

Suppose sensors can be placed at vertices of the graph $\mathcal{G}$; a sensor that is placed at vertex $i \in \mathcal{V}$ gives a measurement of the form

$$(y_k)_i = C_i x_k + (v_k)_i,$$

where $C_i = \mathbf{e}_i^T$ is the state measurement matrix for sensor i, and $(v_k)_i \in \mathbb{R}$ is a zero-mean white Gaussian noise process.

We further define $y_k \triangleq \begin{bmatrix} (y_k)_1 & \cdots & (y_k)_n \end{bmatrix}^T$, $C \triangleq \begin{bmatrix} C_1^T & \cdots & C_n^T \end{bmatrix}^T$ and $v_k \triangleq \begin{bmatrix} (v_k)_1 & \cdots & (v_k)_n \end{bmatrix}^T$. Thus, the output provided by all sensors together is given by

$$y_k = Cx_k + v_k, \tag{2}$$

where $C = I_n$. We denote $\mathbb{E}[v_k v_k^T] = V$ and consider $\mathbb{E}[v_k w_j^T] = \mathbf{0}, \forall k, j \in \mathbb{N}$. The initial state x_0 is also assumed to be independent of v_k for all $k \geq 0$.

A sensor placed at vertex i has a cost $h_i \in \mathbb{R}_{\geq 0}$; define the cost vector $h \triangleq \begin{bmatrix} h_1 & \cdots & h_n \end{bmatrix}^T$. The designer has a budget $H \in \mathbb{R}_{\geq 0}$ that can be spent on placing sensors at the vertices of $\mathcal{G}$.

After the sensors are placed, the Kalman filter is then applied to provide an optimal estimate of the states using the measurements from the installed sensors in the sense of minimizing the mean square estimation error (MSEE). We define a vector $\mu \in \{0, 1\}^n$ as the indicator vector indicating the vertices where sensors are placed. Specifically, $\mu_i = 1$ if and only if a sensor is placed at vertex $i \in \mathcal{V}$. Denote $C(\mu)$ as the measurement matrix of the installed sensors indicated by μ, i.e., $C(\mu) \triangleq \begin{bmatrix} C_{i_1}^T & \cdots & C_{i_p}^T \end{bmatrix}^T$, where $\mathrm{supp}(\mu) = \{i_1, \ldots, i_p\} \subseteq \mathcal{V}$. Similarly, denote $V(\mu)$ as the measurement noise covariance matrix of the installed sensors, i.e., $V(\mu) = \mathbb{E}[v_k(\mu)(v_k(\mu))^T]$, where $v_k(\mu) = \begin{bmatrix} (v_k)_{i_1} & \cdots & (v_k)_{i_p} \end{bmatrix}^T$. Denote $y_{k_1:k_2}$ as the measurements from time step k_1 to k_2, i.e., $y_{k_1:k_2} \triangleq [(y_{k_1})^T \ \cdots \ (y_{k_2})^T]^T$. Denote the (one-step prediction) Kalman filter as $x_{k+1/k}(\mu) = \mathbb{E}[x_{k+1} \big| y_{0:k}(\mu)]$, where $y_{0:k}(\mu) = [(y_0(\mu))^T \ \cdots \ (y_k(\mu))^T]^T$ and $y_i(\mu) \triangleq C(\mu)x_i + v_i(\mu)$ for $i = 0, \ldots, k$. Denote $\tilde{x}_{k+1}(\mu) \triangleq x_{k+1/k}(\mu) - x_{k+1}$ as the estimation error. The *a priori* error covariance matrix of the Kalman filter at time step k, when the sensors indicated by μ are placed and installed, is then given by $\Sigma_{k+1/k}(\mu) \triangleq \mathbb{E}[\tilde{x}_{k+1}(\mu)(\tilde{x}_{k+1}(\mu))^T]$. The limit $\Sigma(\mu) \triangleq \lim_{k \to \infty} \Sigma_{k+1/k}(\mu)$, if it exists, satisfies the *discrete algebraic Riccati equation (DARE)* [17]

$$\Sigma(\mu) = A\Sigma(\mu)A^T + BWB^T - \\ A\Sigma(\mu)C(\mu)^T \big(C(\mu)\Sigma(\mu)C(\mu)^T + V(\mu)\big)^{-1} C(\mu)\Sigma(\mu)A^T. \tag{3}$$

We have the following results from [17].

Lemma 1: Suppose that X, $Y_1, \ldots, Y_k$ are jointly Gaussian random variables. The estimator $\hat{X} = \mathbb{E}[X \big| Y_1, \ldots, Y_k]$ of X given $Y_1, \ldots, Y_k$ is a linear combination of $Y_1, \ldots, Y_k$ and some constants, denoted as $\hat{X} \in \mathcal{L}\{Y_1, \ldots, Y_k\}$. Furthermore, it is also the minimum variance estimator that minimizes $\mathbb{E}[||X - \hat{X}||^2]$ and $\mathbb{E}[(X - \hat{X})(X - \hat{X})^T]$. □

Lemma 2: For a given indicator vector μ, $\Sigma_{k+1/k}(\mu)$ will converge to an unique finite limit $\Sigma(\mu)$ as $k \to \infty$, if and only if the pair $(A, C(\mu))$ is detectable and the pair $(A, BW^{\frac{1}{2}})$ is stabilizable. □

Remark 2: By our assumptions on x_0, w_k and v_k, the Kalman filter is a *linear* minimum variance estimator for all k. Hence, $x_{k+1/k}(\mu) \in \mathcal{L}\{y_0(\mu), \ldots, y_k(\mu)\}$ for all k. □

When the pair $(A, C(\mu))$ is not detectable, we define the limit $\Sigma(\mu) = +\infty$. The Graph-based Kalman Filter Sensor Placement (GKFSP) problem is defined as follows.

Problem 1: (GKFSP) Given a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, consider the matrix $A \in \mathbb{R}^{n \times n}$ as described in Definition 1, a set $\mathcal{I} \subseteq \mathcal{V}$ containing the vertices that have Gaussian inputs, a measurement matrix $C = I_n$ containing all of the individual sensor measurement matrices, a system noise covariance matrix $W \in \mathbb{S}_+^{|\mathcal{I}|}$, a sensor noise covariance matrix $V \in \mathbb{S}_+^n$, a cost vector $h \in \mathbb{R}_{\geq 0}^n$ and a budget $H \in \mathbb{R}_{\geq 0}$, the Graph-based Kalman Filtering Sensor Placement problem is to find the sensor placement μ, i.e., the indicator vector μ of the vertices where sensors are placed, that solves

$$\min_{\mu \in \{0,1\}^n} \quad \text{trace}(\Sigma(\mu))$$
$$s.t.\ h^T \mu \leq H$$

where $\Sigma(\mu)$ is given by Eq. (3) if the pair $(A, C(\mu))$ is detectable, and $\Sigma(\mu) = +\infty$, otherwise.

III. GKFSP: Single input case

In this section, we provide a strategy for the GKFSP problem when there is a single vertex $i_0 \in \mathcal{V}$ that has a Gaussian input, i.e., $\mathcal{I} = \{i_0\} \subseteq \mathcal{V}$, and the graph is a tree.

Given a tree $\mathcal{T} = (\mathcal{V}, \mathcal{E})$ with $\mathcal{V} = \{1, \dots, n\}$, denote the vertex that has the input as $\mathcal{I} = \{i_0\} \subseteq \mathcal{V}$, i.e., $B = \mathbf{e}_{i_0}$. The system dynamics can be written as Eq. (1), where $w_k \in \mathbb{R}$ is a zero-mean white Gaussian noise process with variance denoted as w^2, $\forall k \geq 0$. We will use the following result from [14].

Lemma 3: Given a connected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, consider the matrix A as described in Definition 1 and a discrete-time linear system as defined in Eq. (1) and Eq. (2). Then, the pair (A, Bw) is stabilizable, and the pair $(A, C(\mu))$ is detectable for all sensor placements $\mu \in \{0,1\}^n$ such that $\mu \neq 0$. □

Hence, we know from Lemma 2 and Lemma 3 that if there is a single vertex in the graph that has an input, then $\Sigma_{k+1/k}(\mu)$ converges to the finite limit $\Sigma(\mu)$ as $k \to \infty$, for all $\mu \in \{0,1\}^n$ with $\mu \neq 0$.

We then state the main result of this section.

Theorem 1: Given a tree $\mathcal{T} = (\mathcal{V}, \mathcal{E})$, consider the matrix A as described in Definition 1 and denote the vertex that has the Gaussian input as $\mathcal{I} = \{i_0\} \subseteq \mathcal{V}$. Consider a discrete-time linear system as defined in Eq. (1) and Eq. (2) with $V = \mathbf{0}_{n \times n}$. Then for all sensor placements $\mu \in \{0,1\}^n$ with $\mu \neq 0$, denote $\alpha = \min_{q \in \text{supp}(\mu)} l_{i_0 q}$ to be the shortest distance from vertex i_0 to a sensor indicated by μ. The steady state error covariance matrix $\Sigma(\mu)$ of the Kalman filter satisfies

$$\Sigma(\mu) = w^2 \sum_{j=0}^{\alpha} A^j B B^T A^j.$$

□

Before we prove Theorem 1, we first give some preliminary discussions and results. Consider the system as described in Theorem 1. The response of this system over $\beta + 1$ time steps ($\beta \in \mathbb{Z}_{\geq 0}$) is given by

$$y_{k:k+\beta} = \Theta_\beta x_k + M_\beta w_{k:k+\beta},$$

where $\Theta_\beta = \begin{bmatrix} \Theta_{\beta-1} \\ CA^\beta \end{bmatrix}$ with $\Theta_0 = C$, and $M_\beta = \begin{bmatrix} 0 & 0 \\ \Theta_{\beta-1}B & M_{\beta-1} \end{bmatrix}$ with $M_0 = 0$, The system it is said to have an unknown input state observer with delay β if given the measurements $y_{0:k+\beta}$, the observer produces an estimate $\hat{x}_k$ of the state x_k that converges to x_k as $k \to \infty$, regardless of the value of the input. Once a state observer with delay β is constructed, one can obtain an input observer with delay $\beta + 1$ [18]. We have the following results from [19].

Lemma 4: Consider a discrete-time linear system as defined in Eq. (1) and Eq. (2) with $B = \mathbf{e}_{i_0}$ and $V = \mathbf{0}_{n \times n}$. The system has an unknown input state observer with delay β if and only if

(a) rank$\begin{bmatrix} zI - A & B \\ C & 0 \end{bmatrix} = n + 1$, $\forall z \in \mathbb{C}$, $|z| \geq 1$;

(b) rank$[M_{\beta+1}]-$rank$[M_\beta] = 1$.

□

Remark 3: The estimation error $e_k \triangleq \hat{x}_k - x_k$ of the unknown input state observer as described in [19] satisfies $e_k = E^k e_0$, where E is a stable matrix determined by the system parameters. The initial condition of the observer is $\hat{x}_0$, which does not affect the asymptotic behavior of the observer. Thus, we set $\hat{x}_0 = 0$. Taking the limit as $k \to \infty$, we have $\lim_{k \to \infty} e_k = 0$. Moreover, denoting the state of the input observer as $\hat{w}_k$ and $e'_k \triangleq \hat{w}_k - w_k$, we have $e'_k = \mathbf{e}_{i_0}^T (E - A) E^k e_0$ and $\lim_{k \to \infty} e'_k = 0$. □

We now analyze the conditions in Lemma 4 for the systems that we consider here.

Definition 2: An induced subgraph of $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with vertex set $\mathcal{V}' \subseteq \mathcal{V}$ is the subgraph of $\mathcal{G}$ on $\mathcal{V}'$ containing all edges between those vertices in $\mathcal{G}$. The induced subgraph is denoted as $\mathcal{G}[\mathcal{V}']$. □

Definition 3: Given a connected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, a cut vertex $i \in \mathcal{V}$ is a vertex such that the induced subgraph $\mathcal{G}[\mathcal{V} \setminus \{i\}]$ is not connected. □

Lemma 5: Given a tree $\mathcal{T} = (\mathcal{V}, \mathcal{E})$ and the input vertex set $\mathcal{I} = \{i_0\} \subseteq \mathcal{V}$, consider the matrix A as described in Definition 1 and a discrete-time linear system as defined in Eq. (1) and Eq. (2) with $V = \mathbf{0}_{n \times n}$. The rank condition

$$\text{rank} \begin{bmatrix} zI - A & B \\ C(\mu) & 0 \end{bmatrix} = n + 1, \forall z \in \mathbb{C}, |z| \geq 1 \tag{4}$$

holds for all sensor placements $\mu \in \{0,1\}^n$ with $\mu \neq 0$. □

Proof: We first prove that condition (4) holds for all μ with $|\text{supp}(\mu)| = 1$, i.e., condition (4) holds when $C(\mu) = \mathbf{e}_j^T$, $\forall j \in \mathcal{V}$. We will do it by an induction on $l_{i_0 j}$, i.e., the length of the shortest path between the input vertex i_0 and the output vertex j in $\mathcal{T}$.

First, consider the base case $l_{i_0 j} = 0$, i.e., the sensor is placed at the input vertex i_0. We have rank $\begin{bmatrix} zI - A & \mathbf{e}_{i_0}^T \\ \mathbf{e}_{i_0} & 0 \end{bmatrix} =$ rank$(zI - A_{(i_0,i_0)}) + 2$, where $A_{(i_0,i_0)}$, obtained by removing the i_0th row and i_0th column from A, is a substochastic matrix [16] and is known to be Schur stable [14]. Hence, rank$(zI - A_{(i_0,i_0)}) = n - 1$, $\forall z \in \mathbb{C}, |z| \geq 1$, which implies condition (4) holds when $C(\mu) = \mathbf{e}_{i_0}^T$.

Then, we assume that condition (4) holds when $C(\mu) = \mathbf{e}_j^T$ for $l_{i_0 j} = 1, \dots, \gamma$. Considering $l_{i_0 j} = \gamma + 1$, we can assume without loss of generality that the output vertex $j = 1$, otherwise we can simply relabel the vertices. We then have that

$\text{rank}\begin{bmatrix} zI-A & \mathbf{e}_{i_0} \\ \mathbf{e}_1^T & 0 \end{bmatrix} = \text{rank}\begin{bmatrix} a_1 & 0 \\ zI-A_{(1,1)} & \mathbf{e}_{i_0-1} \end{bmatrix} + 1$, where $a_1 = \begin{bmatrix} A_{12} & A_{13} & \cdots & A_{1n} \end{bmatrix} \in \mathbb{R}^{1\times(n-1)}$. Note that $A_{(1,1)} \in \mathbb{R}^{(n-1)\times(n-1)}$ is the matrix associated with the induced subgraph $\mathcal{T}[\mathcal{V}\setminus\{1\}]$ as described in Definition 1. Since the graph $\mathcal{T}$ is a tree, every vertex in $\mathcal{T}$ that is not a leaf is a cut vertex. Firstly, if vertex 1 is a leaf vertex, we can assume without loss of generality that its only neighbor in $\mathcal{T}$ is vertex 2, which implies $a_1 = \begin{bmatrix} A_{12} & 0 & \cdots & 0 \end{bmatrix}$. Moreover, since the length of the shortest path between i_0 and 1 is $\gamma+1$ in $\mathcal{T}$, and 1 and 2 are connected via edge $(1,2)$ in $\mathcal{T}$, we obtain that the length of the shortest path between i_0 and 2 is γ in $\mathcal{T}[\mathcal{V}\setminus\{1\}]$. Then, by the induction hypothesis, we have $\text{rank}\begin{bmatrix} a_1 & 0 \\ zI-A_{(1,1)} & \mathbf{e}_{i_0-1} \end{bmatrix} = \text{rank}\begin{bmatrix} zI-A_{(1,1)} & \mathbf{e}_{i_0-1} \\ \mathbf{e}_1^T & 0 \end{bmatrix} = n$, $\forall z \in \mathbb{C}$, $|z| \geq 1$. Hence, we have that condition Eq. (4) holds when $C(\mu) = \mathbf{e}_j^T$, where j is a leaf with $l_{i_0 j} = \gamma+1$. Secondly, if vertex 1 is not a leaf, we have that $\mathcal{T}[\mathcal{V}\setminus\{1\}]$ is not connected and has several connected components, denoted as $\mathcal{T}_1', \ldots, \mathcal{T}_r'$, where $r \in \mathbb{Z}_{\geq 2}$. Note that $\mathcal{T}_1', \ldots, \mathcal{T}_r'$ are induced subgraphs of $\mathcal{T}$. Hence, the matrix $A_{(1,1)}$ is a block diagonal matrix of the form $A_{(1,1)} = \text{blkdiag}(A_1', \cdots, A_r')$, where $A_1', \ldots, A_r'$ are associated with $\mathcal{T}_1', \ldots, \mathcal{T}_r'$, respectively. Again, using the fact that $\mathcal{T}$ is a tree, there exists a single vertex in $\mathcal{T}_i'$ that is connected to vertex 1 via an edge in $\mathcal{T}$ for $i = 1, \ldots, r$. It then follows that $A_1', \ldots, A_r'$ are all substochastic matrices, and thus are stable. Furthermore, we assume without loss of generality that the input vertex i_0 is in $\mathcal{T}_1'$ and denote i_1+1 as the vertex in $\mathcal{T}_1'$ that is connected to 1 via edge $(1, i_1+1)$ in $\mathcal{T}$. Hence, the i_1th element of a_1 is the only nonzero element of a_1 from A_{12} to $A_{1(n_1+1)}$, where n_1 is the dimension of A_1'. We then have $\text{rank}\begin{bmatrix} a_1 & 0 \\ zI-A_{(1,1)} & \mathbf{e}_{i_0-1} \end{bmatrix} = \text{rank}\begin{bmatrix} a_1' & a_1'' & 0 \\ zI-A_1' & \mathbf{0} & \mathbf{e}_{i_0-1} \\ \mathbf{0} & zI-A'' & \mathbf{0} \end{bmatrix}$, where $A'' = \text{blkdiag}(A_2', \cdots, A_r')$ is a stable matrix and $a_1 = \begin{bmatrix} a_1' & a_1'' \end{bmatrix}$ with the i_1th element of a_1' to be the only nonzero element of a_1'. Again, since A_1' and A'' are substochastic matrices, we have $\text{rank}\begin{bmatrix} a_1 & 0 \\ zI-A_{(1,1)} & \mathbf{e}_{i_0-1} \end{bmatrix} = \text{rank}\begin{bmatrix} zI-A_1' & \mathbf{e}_{i_0-1} \\ \mathbf{e}_{i_1}^T & 0 \end{bmatrix} + n - 1 - n_1, \forall z \in \mathbb{C}, |z| \geq 1$. Observe that the matrix $\begin{bmatrix} zI-A_1' & \mathbf{e}_{i_0-1} \\ \mathbf{e}_{i_1}^T & 0 \end{bmatrix}$ can be viewed as the matrix pencil when the graph is given by $\mathcal{T}_1'$ with input vertex i_0 (corresponding to $\mathbf{e}_{i_0-1}$) and output vertex i_1+1 (corresponding to $\mathbf{e}_{i_1}^T$). Since the length of the shortest path between 1 and i_0 is $\gamma+1$ in $\mathcal{T}$ and i_1+1 is connected with 1 in $\mathcal{T}$, it follows that the length of the shortest path between i_0 and i_1+1 in $\mathcal{T}_1'$ is γ. By the induction hypothesis, we have that $\text{rank} = \begin{bmatrix} zI-A_1' & \mathbf{e}_{i_0-1} \\ \mathbf{e}_{i_1}^T & 0 \end{bmatrix} = n_1+1$, $\forall z \in \mathbb{C}, |z| \geq 1$. Hence, we have shown that condition (4) holds when $C(\mu) = \mathbf{e}_j^T$, where vertex j is not a leaf with $l_{i_0 j} = \gamma+1$, completing the induction step.

This proves that condition (4) holds when $C(\mu) = \mathbf{e}_j^T$, $\forall j \in \mathcal{V}$. Since adding more rows to a matrix will not decrease its rank, it follows directly that condition (4) is satisfied for all $\mu \in \{0,1\}^n$ with $\mu \neq 0$. ■

Remark 4: Our results here complement the results in [14], where the authors considered the case where the output vertices can also measure the states of their neighbors. Under this assumption, sufficient conditions were provided for condition (4) to hold. However, these sufficient conditions are strong constraints on the locations of the input and output vertices in the graph, i.e., they must be neighbors. We show here that for the single input case, at least one output vertex (without further constraints) is sufficient for condition (4) to hold when the graph is a tree. □

We next analyze condition (b) of Lemma 4. Note that it was shown in [20] that condition (a) in Lemma 4 is sufficient for condition (b) to hold for some β, where β was shown to satisfy $\beta \leq n$ in [21]. We will find the minimum value of β such that condition (b) is satisfied.

Lemma 6: Suppose we are given a connected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and the input vertex $\mathcal{I} = \{i_0\} \subseteq \mathcal{V}$. Consider the matrix A as described in Definition 1 and a discrete-time linear system as defined in Eq. (1) and Eq. (2) with $V = \mathbf{0}_{n\times n}$. For all sensor placements $\mu \in \{0,1\}^n$ with $\mu \neq 0$, denote $\alpha = \min_{q\in\text{supp}(\mu)} l_{i_0 q}$. Then α is the minimum value of β, denoted as β_m, that satisfies the rank condition

$$\text{rank}[M_{\beta+1}] - \text{rank}[M_\beta] = 1.$$

□

Proof: Based on the definition of the matrix A, it is easy to prove by induction the facts that $(A^l)_{ij} = 0$ for all $l < l_{ij}$, where $l \in \mathbb{Z}_{\geq 1}$, and $(A^{l_{ij}})_{ij} > 0$. Hence, we have $C(\mu)A^\beta B = 0$ for all $\beta \leq \alpha - 1$ and $C(\mu)A^\beta B \neq 0$ when $\beta = \alpha$. We then know from the form of M_β that $\beta_m = \alpha$. ■

Remark 5: We know from the above discussions that for the systems that we consider here and a sensor placement $\mu \in \{0,1\}^n$ with $\mu \neq 0$ and $\alpha = \min_{q\in\text{supp}(\mu)} l_{i_0 q}$, there exists an (unknown input) state observer with delay α and an input observer with delay $\alpha+1$. □

We are now in place to prove Theorem 1.

Proof of Theorem 1:

Consider any sensor placement $\mu \in \{0,1\}^n$ with $\mu \neq 0$. For any $k \geq \alpha+1$, we can rewrite x_{k+1} in the form

$$x_{k+1} = A^{\alpha+2} x_{k-\alpha-1} + \sum_{j=0}^{\alpha+1} A^j B w_{k-j}.$$

Moreover, we have

$$\begin{aligned} x_{k+1/k}(\mu) &= \mathbb{E}[A^{\alpha+2} x_{k-\alpha-1} | y_{0:k}(\mu)] \\ &+ \mathbb{E}[A^{\alpha+1} B w_{k-\alpha-1} | y_{0:k}(\mu)] + \mathbb{E}[\sum_{j=0}^{\alpha} A^j B w_{k-j} | y_{0:k}(\mu)]. \end{aligned} \tag{5}$$

Consider the first term on the right hand side of Eq. (5). Denote $x_{k-\alpha-1/k}(\mu) \triangleq \mathbb{E}[x_{k-\alpha-1}|y_{0:k}(\mu)]$. Since $x_{k-\alpha-1}$ and $y_{0:k}(\mu)$ are jointly Gaussian, we know from Lemma 1 that $x_{k-\alpha-1/k}(\mu) \in \mathcal{L}\{y_0, \dots, y_k\}$ for all k. Moreover, we know from Lemma 4 that given the measurements $y_{0:k}(\mu)$, there exists a state observer with delay α (thus, there also exists a state observer with delay $\alpha + 1$) as described in Remark 3 such that $e_{k-\alpha-1} = E^{k-\alpha-1}e_0$, where $e_{k-\alpha-1} = \hat{x}_{k-\alpha-1} - x_{k-\alpha-1}$ and $\hat{x}_{k-\alpha-1}$ is an estimate of $x_{k-\alpha-1}$ given by the observer. We then have $e_{k-\alpha-1}e_{k-\alpha-1}^T = E^{k-\alpha-1}e_0e_0^T(E^T)^{k-\alpha-1}$, which holds everywhere on the sample space of the random variable $e_{k-\alpha-1}$. Taking the expectation on both sides of the equation, we obtain $\mathbb{E}[e_{k-\alpha-1}e_{k-\alpha-1}^T] = E^{k-\alpha-1}\Pi_0(E^T)^{k-\alpha-1}$, where Π_0 is the covariance of the initial condition x_0. Since E is stable, we have $\lim_{k\to\infty}\mathbb{E}[e_{k-\alpha-1}e_{k-\alpha-1}^T] = 0$. Moreover, we know from Remark 2 that $x_{k-\alpha-1/k}(\mu)$ is the minimum variance estimator of $x_{k-\alpha-1}$ for all k. Hence, we have $\lim_{k\to\infty}\mathbb{E}[\tilde{x}_{k-\alpha-1}(\mu)(\tilde{x}_{k-\alpha-1}(\mu))^T] = 0$, where $\tilde{x}_{k-\alpha-1}(\mu) \triangleq x_{k-\alpha-1/k}(\mu) - x_{k-\alpha-1}$. Then, consider the second term on the right hand side of Eq. (5). Using similar arguments as above, we have $\lim_{k\to\infty}\mathbb{E}[(\tilde{w}_{k-\alpha-1}(\mu))^2] = 0$, where $\tilde{w}_{k-\alpha-1}(\mu) \triangleq w_{k-\alpha-1/k}(\mu) - w_{k-\alpha-1}$ and $w_{k-\alpha-1/k}(\mu) \triangleq \mathbb{E}[w_{k-\alpha-1}|y_{0:k}(\mu)]$. Finally, consider the third term on the right hand side of Eq. (5). Since x_0 is assumed to be independent of w_k for all $k \geq 0$, $y_j(\mu)$ is also independent of w_i for $j \leq i$ [22]. We then have

$$\mathbb{E}[\sum_{j=0}^{\alpha} A^j B w_{k-j}|y_{0:k}(\mu)] = \mathbb{E}[\sum_{j=0}^{\alpha} A^j B w_{k-j}|y_{k-\alpha+1:k}(\mu)]. \tag{6}$$

Note that we can rewrite y_k in the form

$$y_k(\mu) = C(\mu)A^\alpha x_{k-\alpha} + \sum_{j=1}^{\alpha} C(\mu)A^{j-1}Bw_{k-j},$$

where $C(\mu)A^{j-1}B = 0$, $\forall j \in \{1, \dots, \alpha\}$. Hence, $y_k(\mu)$ is independent of w_j, $\forall j \in \{k-\alpha, \dots, k\}$. Similarly, $y_{k-1}(\mu)$ is independent of w_j, $\forall j \in \{k-\alpha-1, \dots, k-1\}$. Proceeding in this way, we obtain the following

$$\mathbb{E}[\sum_{j=0}^{\alpha} A^j B w_{k-j}|y_{k-\alpha+1:k}(\mu)] = \mathbb{E}[\sum_{j=0}^{\alpha} A^j B w_{k-j}] = 0. \tag{7}$$

Combining the results above, we have

$$\begin{aligned}\tilde{x}_{k+1}(\mu) = A^{\alpha+2}\tilde{x}_{k-\alpha-1}(\mu) + A^{\alpha+1}B\tilde{w}_{k-\alpha-1}(\mu) \\ + \sum_{j=0}^{\alpha} A^j B w_{k-j}.\end{aligned}$$

Moreover, we have $\tilde{x}_{k-\alpha-1}(\mu) \in \mathcal{L}\{x_{k-\alpha-1}, y_{0:k}\}$ and $\tilde{w}_{k-\alpha-1}(\mu) \in \mathcal{L}\{w_{k-\alpha-1}, y_{0:k}\}$, $\forall k$. We then know from the discussion above that $\tilde{x}_{k-\alpha-1/k}(\mu)$ and $\tilde{w}_{k-\alpha-1}(\mu)$ are both independent of w_j, $\forall j \in \{k-\alpha, \dots, k\}$, $\forall k$. Denoting $\Sigma \triangleq \Sigma_{k+1/k}(\mu) = \mathbb{E}[\tilde{x}_{k+1}(\mu)(\tilde{x}_{k+1}(\mu))^T]$, we have

$$\begin{aligned}\Sigma &= A^{\alpha+2}\Sigma_{k-\alpha-1}A^{\alpha+2} + A^{\alpha+1}B\sigma_{k-\alpha-1}B^TA^{\alpha+1} \\ &+ A^{\alpha+2}\Sigma'_{k-\alpha-1}B^TA^{\alpha+1} + A^{\alpha+1}B\Sigma'^T_{k-\alpha-1}A^{\alpha+2} \\ &+ w^2\sum_{j=0}^{\alpha} A^jBB^TA^j\end{aligned}$$

where $\Sigma_{k-\alpha-1} \triangleq \mathbb{E}[\tilde{x}_{k-\alpha-1}(\mu)(\tilde{x}_{k-\alpha-1}(\mu))^T]$, $\sigma_{k-\alpha-1} \triangleq \mathbb{E}[(\tilde{w}_{k-\alpha-1}(\mu))^2]$ and $\Sigma'_{k-\alpha-1} \triangleq \mathbb{E}[\tilde{x}_{k-\alpha-1}(\mu)(\tilde{w}_{k-\alpha-1}(\mu))]$. Note that we have shown that $\lim_{k\to\infty}\Sigma_{k-\alpha-1} = 0$ and $\lim_{k\to\infty}\sigma_{k-\alpha-1} = 0$. Considering $\Sigma'_{k-\alpha-1}$ element-wise, we have $(\mathbb{E}[(\tilde{x}_{k-\alpha-1}(\mu))_i\tilde{w}_{k-\alpha-1}(\mu)])^2 \leq \mathbb{E}[(\tilde{x}_{k-\alpha-1}(\mu))_i^2]\mathbb{E}[(\tilde{w}_{k-\alpha-1}(\mu))^2]$, $\forall i$, by the Cauchy-Schwarz inequality. Since $\lim_{k\to\infty}\mathbb{E}[(\tilde{x}_{k-\alpha-1}(\mu))_i^2] = 0$, $\forall i$, and $\lim_{k\to\infty}\mathbb{E}[(\tilde{w}_{k-\alpha-1}(\mu))^2] = 0$, we have $\lim_{k\to\infty}\mathbb{E}[(\tilde{x}_{k-\alpha-1}(\mu))_i w_{k-\alpha-1}(\mu)] = 0$, $\forall i$. Combining the results above, we obtain

$$\Sigma(\mu) = \lim_{k\to\infty}\Sigma_{k+1/k}(\mu) = w^2\sum_{j=0}^{\alpha} A^jBB^TA^j. \tag{8}$$

This completes the proof of Theorem 1. ∎

Remark 6: Note that the result in Eq. (8) can be verified by directly plugging it into the DARE as defined in Eq. (3). However, our analysis here sheds light on the relationship between the delayed unknown input observer observer and the Kalman filter, providing an alternative way to calculate the steady-state error covariance of the Kalman filter. □

Remark 7: Note that from Eq. (8), given the matrices A and B, the value of the error covariance matrix only depends on α, i.e., the shortest distance from the sensors to the input. Hence, it is enough to consider sensor placements $\mu \in \{0,1\}^n$ such that $|\text{supp}(\mu)| = 1$, i.e., it is enough to place a single sensor at a single vertex in order to minimize the steady-state MSEE. □

Corollary 1: Given a tree $\mathcal{T} = (\mathcal{V}, \mathcal{E})$, consider the matrix A as described in Definition 1 and the set of input vertex as $\mathcal{I} = \{i_0\} \subseteq \mathcal{V}$. Consider a discrete-time linear system as defined in Eq. (1) and Eq. (2) with $V = \mathbf{0}_{n\times n}$. For any two sensor placements $\mu, \mu' \in \{0,1\}^n$ and $\mu, \mu' \neq 0$ with $\text{supp}(\mu) = \{j\}$, $\text{supp}(\mu') = \{j'\}$, the following results hold.

(a) For all $r \in \mathcal{V}$, $(\Sigma(\mu))_{rr} \geq (\Sigma(\mu'))_{rr}$ if and only if $l_{i_0j} \geq l_{i_0j'}$.
(b) For all $r \in \mathcal{V}$ such that $l_{ir} > l_{i_0j}$, $(\Sigma(\mu))_{rr} = 0$.

□

Proof: (a). Since $A^jBB^TA^j$ is a nonnegative matrix for $j \in \mathbb{Z}_{\geq 0}$, the result follows directly from Eq. (8).

(b). For all $r \in \mathcal{V}$ such that $l_{ri_0} > \alpha$, we know that $(A^j)_{ri_0} = 0$ for all $j = 0, 1, \dots, \alpha$, which implies $(A^jBB^TA^j)_{rr} = 0$ for all $j = 0, 1, \dots, \alpha$. Hence, we have $(\Sigma(\mu))_{rr} = 0$. ∎

Using the results above, the optimal solution to the GKFSP problem as defined in Problem 1, when the graph is a tree with $\mathcal{I} = \{i_0\} \subseteq \mathcal{V}$, is to put a sensor at a vertex $j \in \mathcal{V}$ that is as close as possible to i_0 with respect to the distance l_{i_0j} in the tree $\mathcal{T}$, while satisfying the budget constraint. Thus, polynomial-time algorithms such as breadth-first-search (BFS) can be used [23].

IV. GKFSP: Multi-input case

In this section, we analyze the case for GKFSP when there are multiple vertices in the graph that have Gaussian inputs. Denote the set of vertices that have inputs as $\mathcal{I} = \{i_0, \ldots, i_{\tau-1}\} \subseteq \mathcal{V}$, where $\tau \in \mathbb{Z}_{\geq 2}$.

Theorem 2: The GKFSP problem is NP-hard when the system has multiple inputs. □

Proof: We prove by giving a reduction from the knapsack problem [24] to GKFSP when there are multiple inputs in the graph. Consider an arbitrary instance of the knapsack problem to be the number of items q, the set of values $\{\gamma_i\}$, the set of weights $\{\kappa_i\}$, the weight budget $\bar{\kappa}$, and an indicator vector $\nu \in \{0,1\}^q$ such that item i is chosen if and only if $\nu_i = 1$, where $\{\gamma_i\}$, $\{\kappa_i\}$ and $\bar{\kappa}$ are positive integers. We construct an instance of GKFSP with multiple inputs as follows. The graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is set to consist of q vertex-disjoint paths, all of which are of length 2. Denote $\mathcal{G} = (\mathcal{P}_1, \ldots, \mathcal{P}_q)$, where the vertex set of $\mathcal{P}_i$ is denoted as $\mathcal{V}_i = \{p_1^i, p_2^i, p_3^i\}$ for $i = 1, \ldots, q$. The matrix $A \in \mathbb{R}^{3q \times 3q}$ associated to $\mathcal{G}$ is then as described in Definition 1 with $\epsilon = \frac{1}{2d_{max}}$, denoted as $A = \text{blkdiag}(A_1, \ldots, A_q)$, where A_i is the matrix associated to $\mathcal{P}_i$. The measurement matrix is $C = I_{3q}$. The input vertex set is $\mathcal{I} = \{p_1^1, p_1^2, \ldots, p_1^q\}$. The system noise covariance matrix is set as $W_{\mathcal{I}} = \text{diag}(w_1, \ldots, w_q)$ with $w_i = \frac{\gamma_i}{(A_i^4)_{11}}$. The measurement noise covariance matrix is $V = \mathbf{0}_{3q \times 3q}$. The sensor placement budget $H = \bar{\kappa}$. The cost vector $h = [\mathbf{h}_1^T \ \cdots \ \mathbf{h}_q^T]^T$, where $\mathbf{h}_i = [\bar{\kappa}+1 \ \kappa_i \ 0]^T$ is the cost vector for putting sensors at vertices of $\mathcal{P}_i$. Hence, a sensor will always be placed at p_3^i for all i. Note that since A, C, W and V are all (block) diagonal, we have the steady-state error covariance matrix of the Kalman filter, if it exists, satisfies $\Sigma(\mu) = \text{blkdiag}(\Sigma_1(\mu_1) \ldots, \Sigma_q(\mu_q))$, where μ_i is the sensor placement vector w.r.t. $\mathcal{P}_i$ in $\mathcal{G}$ and $\Sigma_i(\mu_i)$ is the associated error covariance matrix. Since each $\mathcal{P}_i$ is a tree, the results in the previous section can be applied here. Specifically, we have $\text{trace}(\Sigma_i(\mu_i)\big|_{\mu_i=[0\ 0\ 1]^T}) - \text{trace}(\Sigma_i(\mu_i)\big|_{\mu_i=[0\ 1\ 1]^T}) = w_i \mathbf{e}_1^T A_i^4 \mathbf{e}_1 = \gamma_i$ for all i. Consider a sensor placement $\tilde{\mu}$ with $\text{supp}(\tilde{\mu}) \subseteq \mathcal{V}$ such that $\{p_2^{j_1}, \ldots, p_2^{j_\sigma}\} \in \text{supp}(\tilde{\mu})$, i.e., the middle points of paths $\mathcal{P}_{j_1}, \ldots, \mathcal{P}_{j_\sigma}$ are chosen to place sensors. Then, by our construction of the GKFSP instance as above, $\tilde{\mu}$ is optimal if and only if the indicator vector ν with $\text{supp}(\nu) = \{j_1, \ldots, j_\sigma\}$ is optimal for the knapsack instance. Since the knapsack problem is NP-hard, the GKFSP problem is NP-hard when the system has multiple inputs. ■

V. Conclusions

In this paper, we studied the GKFSP problem for a class of linear dynamical systems defined over consensus networks with stochastic inputs. We showed that there exist polynomial-time algorithms for this problem when there is a single stochastic input in a tree network. We further showed that the GKFSP problem becomes NP-hard when the system has multiple inputs. Our analysis sheds light on the relationship between the Kalman filter and the (delayed) unknown input observer and provides an explicit solution to the DARE associated with the error covariance matrix of the Kalman filter. Future work on extending the results to more general network topologies and providing approximation algorithms for the multi-input case are of interest.

References

[1] A. Olshevsky, "Minimal controllability problems," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 3, pp. 249–258, 2014.

[2] S. Pequito, S. Kar, and A. P. Aguiar, "Minimum cost input/output design for large-scale linear structural systems," *Automatica*, vol. 68, pp. 384–391, 2016.

[3] D. J. Chmielewski, T. Palmer, and V. Manousiouthakis, "On the theory of optimal sensor placement," *AIChE journal*, vol. 48, no. 5, pp. 1001–1012, 2002.

[4] C. Yang, J. Wu, X. Ren, W. Yang, H. Shi, and L. Shi, "Deterministic sensor selection for centralized state estimation under limited communication resource," *IEEE Transactions on Signal Processing*, vol. 63, no. 9, pp. 2336–2348, 2015.

[5] V. Tzoumas, A. Jadbabaie, and G. J. Pappas, "Sensor placement for optimal Kalman filtering: Fundamental limits, submodularity, and algorithms," in *Proc. American Control Conference*, 2016, pp. 191–196.

[6] H. Zhang, R. Ayoub, and S. Sundaram, "Sensor selection for Kalman filtering of linear dynamical systems: Complexity, limitations and greedy algorithms," *Automatica*, vol. 78, pp. 202–210, 2017.

[7] L. F. Chamon, G. J. Pappas, and A. Ribeiro, "The mean square error in Kalman filtering sensor selection is approximately supermodular," in *Proc. IEEE Conference on Decision and Control*, 2017, pp. 343–350.

[8] L. Ye, S. Roy, and S. Sundaram, "On the complexity and approximability of optimal sensor selection for Kalman filtering," in *Proc. American Control Conference*, 2018, pp. 5049–5054.

[9] K. Fitch and N. E. Leonard, "Joint centrality distinguishes optimal leaders in noisy networks," *IEEE Transactions on Control of Network Systems*, vol. 3, no. 4, pp. 366–378, 2016.

[10] F. Lin, M. Fardad, and M. R. Jovanović, "Algorithms for leader selection in stochastically forced consensus networks," *IEEE Transactions on Automatic Control*, vol. 59, no. 7, pp. 1789–1802, 2014.

[11] M. Pirani, E. M. Shahrivar, B. Fidan, and S. Sundaram, "Robustness of leader-follower networked dynamical systems," *IEEE Transactions on Control of Network Systems*, 2017.

[12] S. Roy, M. Xue, and S. Sundaram, "Graph-theoretic analysis of estimators for stochastically-driven diffusive network processes," in *Proc. American Control Conference*, 2018, pp. 1796–1801.

[13] S. Sundaram and C. N. Hadjicostis, "Distributed function calculation via linear iterative strategies in the presence of malicious agents," *IEEE Trans. on Automatic Control*, vol. 56, no. 7, pp. 1495–1508, 2011.

[14] F. Pasqualetti, A. Bicchi, and F. Bullo, "Consensus computation in unreliable networks: A system theoretic approach," *IEEE Transactions on Automatic Control*, vol. 57, no. 1, pp. 90–104, 2012.

[15] J.-M. Dion, C. Commault, and J. Van Der Woude, "Generic properties and control of linear structured systems: a survey," *Automatica*, vol. 39, no. 7, pp. 1125–1144, 2003.

[16] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[17] B. D. Anderson and J. B. Moore, *Optimal Filtering*. Dover Books, 1979.

[18] J. Jin, M.-J. Tahk, and C. Park, "Time-delayed state and unknown input observation," *International Journal of Control*, vol. 66, no. 5, pp. 733–746, 1997.

[19] S. Sundaram and C. N. Hadjicostis, "Delayed observers for linear systems with unknown inputs," *IEEE Transactions on Automatic Control*, vol. 52, no. 2, pp. 334–339, 2007.

[20] P. Moylan, "Stable inversion of linear systems," *IEEE Transactions on Automatic Control*, vol. 22, no. 1, pp. 74–78, 1977.

[21] A. Willsky, "On the invertibility of linear systems," *IEEE Transactions on Automatic Control*, vol. 19, no. 3, pp. 272–274, 1974.

[22] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear estimation*. Prentice Hall, 2000, no. EPFL-BOOK-233814.

[23] T. H. Cormen, *Introduction to algorithms*. MIT press, 2009.

[24] M. R. Garey and D. S. Johnson, *Computers and intractability: a guide to the theory of NP-completeness*. Freeman, 1979.