ON THE DISTRIBUTED ESTIMATION OF RANK-DEFICIENT DYNAMICAL SYSTEMS: A GENERIC APPROACH

Mohammadreza Doostmohammadian and Usman A. Khan

Department of Electrical and Computer Engineering, Tufts University, Medford, MA, 02155.
{mrd, khan}@ece.tufts.edu

ABSTRACT

In this paper, we consider distributed estimation when the communication time-scale is restricted to the time-scale of the dynamics. It can be shown that this restriction may not guarantee a stable estimation error when the data fusion is implemented only in the *observation-space*. To address this issue, one has to rely on fusion in the *predictor-space*, which *alone* may lead to a stable error *only* when the system matrix is full S-rank (maximal rank of the zero/non-zero structure). In this paper, we show that when the system matrix is S-rank deficient, predictor-space fusion is insufficient, i.e., the distributed estimator is not observable. In order to recover distributed observability, we provide a novel measurement-based agent classification, and subsequently, define inter-agent communication derived from this classification. The results are based on structured systems theory and the notion of generic observability. Finally, we provide an illustrative example to show the applicability of the proposed schemes using an iterative Linear Matrix Inequality (LMI) approach.

Index Terms— Distributed estimation, S-rank, structured systems theory, generic observability, graph theory

1. INTRODUCTION

Distributed estimation is where a network of agents is tasked to estimate the state of a dynamical system when the agents can only communicate over a sparse communication network. Recently [1–4], consensus-based estimation has been proposed as a distributed solution of this problem where the agents implement a message-passing algorithm over the *observation space* between every two successive time-steps, k and $k+1$, of the system dynamics. For optimal performance, this consensus-based estimator requires a consensus to be reached that subsequently results into infinite (or very large) information exchanges when the agent communication is sparse.

When the network is unable to implement a consensus due to, e.g., resource-constraints or faster system dynamics, distributed solutions have been proposed with finite information exchanges [5–11] at the price of requiring an agreement over *predictor-space*. However, in the finite-time scenario, one first has to guarantee the existence of a communication network that will result into a bounded estimation error, i.e., distributed observability. Clearly, the problem of distributed observability is only challenging when none of the agents is observable alone or in its neighborhood, but only the entire network as a whole guarantees observability; this is a typical assumption in related work and also in this paper.

Existing work in this regard is restricted to system matrices that have a full structured-rank (maximal rank of the zero non-zero structure) and strongly-connected communication networks [6–10] at the price of requiring an agreement over *predictor-space*. For example, [6, 7] established results on full S-rank systems where the communication is assumed to be strongly-connected in the former and weakly-connected in the latter. Distributed estimation based on moving horizon estimation [8], information theoretic approach [9], and diffusion-based methods [10] is also proposed in the literature; all of them requiring the system matrix to be invertible (full S-rank). However, in many practical applications the system matrix is S-rank deficient, e.g., detection/estimation problems in [12], Gauss-Markov system models in [13], smart grids [14], and Type-C Wind Turbine Generator models [15].

We use structured systems theory and the notion of generic observability. Generic properties are useful in, e.g., models where the environment uncertainties are reflected in the system parameters. As long as the system structure (zero/non-zero pattern) is not violated, a generic property that holds for one set of parameters also hold for *almost all* choices of non-zero parameters. This leads to a robust estimator design where the analysis is not algebraic (as in the conventional Grammian or PBH rank tests for observability), but graph-theoretic [11, 16–20]. In the graph theoretic approach considered here, we employ concepts such as maximal sets, cycle families, Strongly Connected Component (SCC) to provide a novel agent classification. We classify agents–based on their role in global observability–as crucial and non-crucial; and further, subdivide the crucial agents with respect to their role in recovering *distributed observability*. Subsequently, this classification leads to a network topology design with minimal communication.

This main contributions of this paper include: (i) Only *one* information exchange is allowed among the agents. (ii) For a system matrix with full S-rank, a weakly-connected network is sufficient to guarantee observability; in addition, the requirements on the underlying network topology are explicitly characterized. (iii) When the system matrix is S-rank deficient, no agent communication network can guarantee observability with agreement in the *predictor-space* alone, and hence, fusion in the *observation-space* is required. (iv) The results are structure-based in contrast to widely used algebraic (rank-based) tests for observability. (v) The proposed solutions are shown to work with constrained LMI-based iterative gain matrix design (see [6] for more details on this). The novelty of this work further lies in the fact that widely-used algebraic notions are replaced with structure-based generic notions and has applications in network topology design for smart-grids and multi-agent systems [12–15].

We now describe the rest of the paper. Section 2 provides preliminary material and terminologies. In Section 3, we present the problem formulation along with the assumptions and the proposed agent-classification method. We describe main results on S-rank deficient systems in Section 4, which are further explained via an illustrative example in Section 5. Finally, Section 6 concludes the paper.

2. PRELIMINARIES

Consider a discrete-time linear dynamical system:

$$\mathbf{x}_{k+1} = A\mathbf{x}_k + \mathbf{v}_k, \qquad \mathbf{y}_k = Cx_k + \mathbf{r}_k, \tag{1}$$

where $\mathbf{x}_k \in \mathbb{R}^n$ is the state vector; $A \in \mathbb{R}^{n\times n}$ is the system matrix, and $\mathbf{v}_k \sim \mathbb{N}(0, V)$ is the system noise; C is the global output matrix; $\mathbf{y}_k$ is the collection of local output vectors $\mathbf{y}_k^i \in \mathbb{R}^{m_i}$; $\mathbf{r}_k \sim \mathbb{N}(0, R)$ is the global observation noise with $R = \text{blockdiag}[R_1, \ldots, R_N]$; and $\mathbf{r}_k^i \sim \mathbb{N}(0, R_i)$ is the local output noise at agent i.

Let $\widehat{\mathbf{x}}_{k|k}^c$ be the centralized Kalman estimator [21] at time k given all the observations, $\mathbf{y}_k$, up to time k. The error in the centralized Kalman estimator, $\widehat{\mathbf{e}}_{k|k}^c = \mathbf{x}_k - \widehat{\mathbf{x}}_{k|k}^c$, is given by

$$\widehat{\mathbf{e}}_{k|k}^c = (A - K_c CA)\widehat{\mathbf{e}}_{k-1|k-1}^c + \eta_k, \tag{2}$$

where K_c is the centralized Kalman gain and the vector η_k collects the remaining (noise) terms that are independent of $\widehat{\mathbf{e}}_{k-1|k-1}^c$. It is well known that the centralized Kalman error, $\widehat{\mathbf{e}}_{k|k}^c$, is stable if and only if (A, C) is *generically observable* (to be defined later).

2.1. Graph terminology

Let $X = \{x_1, \ldots, x_n\}$ denote the state set, and $Y = \{y_1, \ldots, y_N\}$ denote the output set. We define two different graph representations: *system digraph*, $G_A = (V, E)$, where $V = X \cup Y$ is the vertex set representing states of the dynamic system, and E defines the edge set representing state interaction. The other digraph, G_W, determines the *communication network* of the agents monitoring system states. Let $G_W = (V_W, E_W)$, where $V_W = \{y_1, \ldots, y_N\}$ (or, simply, $V_W = \{1, \ldots, N\}$) is the set of agents, $E_W = \{(i,j) \mid i \leftarrow j\}$ defines the set of communication links among agents, and $\mathcal{D}_i = \{i\} \cup \{j \mid (i,j) \in E_W\}$ denote the extended neighborhood of agent i. Notice that, in contrast to existing works we do not constrain G_W to be undirected. In fact, no assumption on the topology is considered here, as designing G_W is a contribution of the paper. We omit the standard definitions of cycle, (simple) path, and Strongly Connected Component (SCC) which can be found in [22]. We review some new definitions here (see [7] for more details): in G_A, we define *parent SCC* as an state SCC with no outgoing edge to any other SCC; and we call a non-parent SCC a *child SCC*.

2.2. Notes on Structured systems theory

Here, we present definitions and results on generic rank and generic observability. Due to space limitation, we omit the proofs, and refer interested readers to [11, 17–19].

Definition 1 (S-rank). *S-rank (structural rank or generic rank) of a matrix, A, is the maximal rank over all numerical values of the non-zero entries of the matrix A.*

Lemma 1. *A system matrix, A, is full S-rank if and only if there exists a disjoint family of cycles spanning all the state vertices in G_A; otherwise, the system is S-rank deficient.*

Examples of S-rank deficient systems are shown in Fig. 1. For both graphs, there are no family of *disjoint* cycles spanning all the state nodes.

Theorem 1. *A dynamical system is generically observable if and only if in its system digraph: (i) every state is the begin-node of a path that ends in an output (termed as a Y-topped path); and (ii) there exists a disjoint union of Y-topped paths and cycles that covers all the state vertices.*

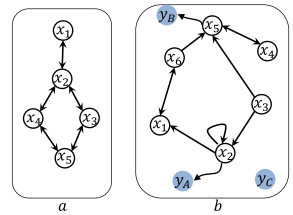

Fig. 1. S-rank deficiency: (a) undirected and (b) directed graph (along with observation vertices).

As an example consider the system shown in Fig. 1(b). It can be verified that each state is a begin-vertex of a Y-topped path, and $\{(x_1, x_6), (x_4, x_5), (x_3, x_2, y_A)\}$ constitute a disjoint family of cycles and Y-topped paths. Thus, the system is *generically* (A, C) observable.

Lemma 2. *The condition (ii) in Theorem 1 on the generic observability of $(A_{n\times n}, C_{N\times n})$ is equivalent to S-rank$([A^T\ C^T]^T) = n$.*

3. DISTRIBUTED KALMAN-TYPE ESTIMATOR

Let $\widehat{x}_{k|m}^i$ be the state estimate at time k and agent i given the outputs up to time m, $(m \leq k)$, from agent i and its neighbors, $j \in \mathcal{D}_i$. Each agent implements the following variant of distributed Kalman-type Estimation (DKE) as proposed in [6, 7]:

$$\widehat{\mathbf{x}}_{k|k-1}^i = \sum_{j\in\mathcal{D}_i} w_{ij} A \widehat{\mathbf{x}}_{k-1|k-1}^j, \tag{3}$$

$$\widehat{\mathbf{x}}_{k|k}^i = \widehat{\mathbf{x}}_{k|k-1}^i + K_k^i \sum_{j\in D_i} C_j^T (y_k^j - C_j \widehat{\mathbf{x}}_{k|k-1}^i), \tag{4}$$

where Eq. (3) is the local predictor representing fusion in the predictor-space and Eq. (4) represents fusion in observation-space. The corresponding variables are: $W = \{w_{ij}\}$ is the fusion weight matrix for predictor-space, such that $w_{ij} \geq 0$ with $\sum_{j\in\mathcal{D}_i} w_{ij} = 1$ (stochastic), and K_k^i is the local gain matrix at agent i and time k.

Let $\mathbf{e}_k^i = \mathbf{x}_k - \mathbf{x}_{k|k}^i$ be the local estimation error at agent i and time k and $\mathbf{e}_k = [(\mathbf{e}_k^1)^T, \ldots, (\mathbf{e}_k^N)^T]^T$ be the network estimation error. After some straightforward manipulations (see [6]), we obtain the following distributed error dynamics:

$$\mathbf{e}_k = (W \otimes A - K_k D_C (W \otimes A))\mathbf{e}_{k-1} + \mathbf{q}_k, \tag{5}$$

where $\mathbf{q}_k = [(\mathbf{q}_k^1)^T, \ldots, (\mathbf{q}_k^N)^T]^T$ collects the remaining terms which are weighted linear function of the system and output noise independent of $\mathbf{e}_k$ and,

$$K_k = \text{blockdiag}[K_k^1, \ldots, K_k^N], \tag{6}$$

$$D_C = \text{blockdiag}[\sum_{j\in\mathcal{D}_1} C_j^T C_j, \ldots, \sum_{j\in\mathcal{D}_N} C_j^T C_j], \tag{7}$$

Comparing Eq. (5) with Eq. (2), it is straightforward to see that the distributed estimation error, $\mathbf{e}_k$, can be stabilized if and only if $(W \otimes A, D_C)$ is Therefore, the communication network, W, plays a major role for error stability of the single time-scale distributed estimation.

Remarks:

(R1) Every block diagonal, $\sum_{j\in\mathcal{D}_i} C_j^T C_j$, in the matrix D_C, can be thought of as a representation of all the observations in the extended neighborhood, $\mathcal{D}_i$, of agent i.

(R2) The diagonal entries of W are all nonzero, since every agent is in its own extended neighborhood. Therefore, G_W has a disjoint union of self-cycles, and S-rank$(W) = N$.

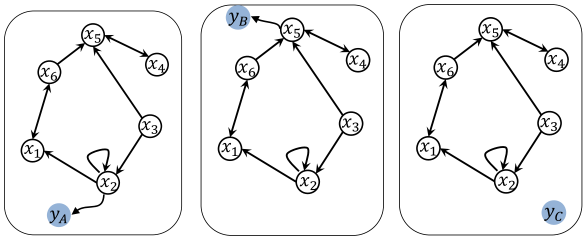

Fig. 2. The graph associated with $(I \otimes A, \overline{D_C})$. Agent y_C has no access to any state of the system. System is (A, C) observable but not (A, C_i) observable for any agent.

3.1. Assumptions

In the rest of the paper, we make the following assumptions:
(i) The communication between the agents is stable, i.e., the communication network is static;
(ii) The system is globally (A, C)-observable;
(iii) For every agent, i, the pairs, (A, C_i) or $(A, \sum_{j \in \mathcal{D}_i} C_j^T C_j)$, are not necessarily observable.
Assumption (ii) is a typical assumption in distributed estimation implying the observability of centralized estimator; without this, no estimation scheme will work. Assumption (iii), in practice, makes the distributed estimation problem more challenging and is where this work becomes significantly different from many current approaches.

3.2. Problem formulation

We now characterize the distributed observability of the estimator in Eqs. (3)-(4), i.e., the observability of $(W \otimes A, D_C)$. We refer to $(W \otimes A, D_C)$ as the *distributed system*, and $G_{W \otimes A}$ as its associated digraph. Consider $W = I$ and $D_C = \overline{D_C}$ where $\overline{D_C} = \text{blockdiag}[C_1^T C_1, \ldots, C_N^T C_N]$; this implies no information exchange among the agents, where the distributed system, $(I \otimes A, \overline{D_C})$, consists of N subsystems associated with each agent (for example, see Fig. 2). With no information fusion, each agent has only partial observation of the system, and has to obtain the missing information from its immediate neighborhood over time. In system digraph this information sharing provides more linking among the subsystems. This extra linking, captured by the non-zeros in W and the summation in D_C, has the potential to improve the generic observability of the system.

In this regard, the distributed observability can be recovered via either $W \otimes A$ matrix (i.e. fusion in predictor-space) or D_C (i.e. fusion in observation-space). In observation-space, a link between two agents, for example from j to i ($j \rightarrow i$), implies that agent i has *only* access to agent j's measurement. Mathematically, for this case, the distributed system is defined as $(I \otimes A, D_C)$. On the other hand, for fusion in predictor-space, the distributed system is $(W \otimes A, \overline{D_C})$. In this more challenging case, adding a link from j to i, reflects as edges from some states in the subsystem of agent j to some states in the subsystem of agent i. This is more discussed in Section 4.

3.3. Agent classification

To describe our approach, we provide a novel agent classification. Since the system is (A, C)-observable (and we assume that an A, C pair is given), condition (i) in Theorem 1 enlists a disjoint union of cycles and Y-topped paths, in the system graph, G_A, that covers all the state vertices. Among possible choices of we find the *maximal* set, $\mathcal{L}$ with the largest number of vertices contained in its cycles/paths. For example, consider the graph in Fig. 1(b) with three agents: y_A observing output from x_2, y_B observing output from x_5, and y_C with no observation. Two options for the maximal set, $\mathcal{L}$, are: $\mathcal{L}_1 = \{(x_1, x_6), (x_4, x_5), (x_3, x_2, y_A)\}$ and $\mathcal{L}_2 = \{(x_1, x_6), (x_4, x_5, y_B), (x_3, x_2, y_A)\}$, among others. We perform the following agent classification with respect to parent/child SCC classification and maximal set, $\mathcal{L}$:

Definition 2. Type-α agent *is the one that appears in the Y-topped paths in $\mathcal{L}$. For example, agent y_A in Fig. 2.* Type-β agent *is the one that measures a state in a full S-rank parent SCC in $\mathcal{L}$. For example, agent y_B in Fig. 2.* Type-γ agent *is the one that is not Type-α or Type-β. For example, agent y_C in Fig. 2.*

Noting that $\{x_4, x_5\}$ is a full S-rank parent SCC, we get $\{y_A, y_B, y_C\}$ as Type-$\{\alpha, \alpha, \gamma\}$ agents for $\mathcal{L}_1$, and Type-$\{\alpha, \beta, \gamma\}$ agents for $\mathcal{L}_2$. We now define *crucial* agents in the sense that removing them renders the system unobservable. This is different from crucial agents for global observability [16]; this is because we further subdivide crucial agents into two categories.

Lemma 3. *The following are crucial for observability: (i) Every Type-α agent; and (ii) At least one Type-β agent observing a state in every full S-rank parent SCC, $\mathcal{K}$, in G_A.*

As a sketch of the proof, note that removing a Type-α agent violates condition (i), while having a parent SCC, $\mathcal{K}$, with no Type-β agent violates condition (ii) in Theorem 1 (see [7] for details.). For example, in Fig. 2, both agent y_A and agent y_B are crucial for observability.

4. RECOVERING DISTRIBUTED OBSERVABILITY

In this section, we present main results on the role of fusion in predictor-space and in observation-space for $(W \otimes A, D_C)$ observability of S-rank deficient system. The proofs are mainly graph theoretic that is a direct consequence of our generic approach. In the following, Lemma 4 follows from Theorem1 while Theorem 2 provides the main result on fusion in predictor-space. Due to space limitation, we omit the proofs of the lemmas and theorems in this section and refer the interested readers to [23].

Lemma 4. *The matrix $W \otimes A$ is S-rank deficient, if and only if the matrix $A_{n \times n}$ is S-rank deficient.*

Proof. The proof directly follows from the definition S-rank$(W \otimes A) = \max(rank(W \otimes A)) = N \times n$. See detailed proof in [23] □

Theorem 2. *If system, A, is S-rank deficient, then $(W \otimes A, \overline{D_C})$ is not observable for any choice of the matrix W.*

Proof. The detailed proof is given in [23]. □

The above theorem shows that for S-rank-deficient systems, fusion in predictor-space does not guarantee distributed observability. Thus, to recover observability, agents need more outputs that implies fusion in observation-space. For fusion in observation-space, the structure of the matrix D_C has to be determined such that $(I \otimes A, D_C)$ is observable. Note that the ith $n \times n$ diagonal block of D_C contains all outputs in the extended neighbourhood $\mathcal{D}_i$. In $G_{(I \otimes A, D_C)}$, this means direct access to outputs in $\mathcal{D}_i$. This follows to our main result on fusion in observation-space.

Theorem 3. *Having the assumptions (i)-(iii) in Section 3.1, the system $(I \otimes A, D_C)$ is observable if and only if in the communication network G_W:*
(i) Every Type-α agent, i, is directly linked to every other agent j;
(ii) For every full S-rank parent SCC, $\mathcal{K}$, every agent, j, without a state observation in $\mathcal{K}$ is directly linked to a Type-β agent, i, with a state observation in $\mathcal{K}$.

Proof. The proof is a direct result of Lemma 3 (see [23]). □

Finally, consolidating this with Theorem 2, we directly get the main theorem on generic observability of S-rank-deficient systems.

Theorem 4. *Let the assumptions (i)-(iii) in Section 3.1 hold. For $(W \otimes A, D_C)$ observability with minimal sufficient communications, each agent needs:*
(i) A direct link from all the Type-α agents;
(ii) A directed path to (at least) one Type-β agent for every full S-rank parent SCC of A. This means, if there are two or more outputs from the same SCC, a directed path to any one of them is sufficient.

Remarks:
(R3) Notice that, every agent requires a directed *path to* each Type-β agent while a direct *link from* each Type-α agent. Therefore, roughly speaking, Type-β agents requires less number of links compared to Type-α agents.
(R4) The generic observability of $(W \otimes A, D_C)$ implies existence of a *full* gain matrix, K, such that $\rho(W \otimes A - K_k D_C(W \otimes A)) < 1$. Here, however, the gain matrix, K_k, is needed to be *block-diagonal* with N blocks of $n \times n$ matrices. To find such matrix, we use the iterative procedure based on LMI approach proposed in [6]. Further, it is assumed that the matrix K_k is independent of time, k. We omit the details here and refer interested readers to [6, 23].

5. EXAMPLE AND SIMULATION

For simulation, reconsider the system structure in Fig. 1(b). The structure of the matrices A and C are as follows:

$$A = \begin{bmatrix} 0 & \times & 0 & 0 & 0 & \times \\ 0 & \times & \times & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \times & 0 \\ 0 & 0 & \times & \times & 0 & 0 \\ \times & 0 & 0 & 0 & \times & 0 \end{bmatrix}, \quad (8)$$

$$C = \begin{bmatrix} y_A \\ y_B \\ y_C \end{bmatrix} = \begin{bmatrix} 0 & \times & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \times & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \quad (9)$$

Based on Theorem 4, we propose communication matrices W_1 and W_2, respectively, for maximal sets $\mathcal{L}_1$ and $\mathcal{L}_2$ (given in section 3.3):

$$W_1 = \begin{bmatrix} \times & \times & 0 \\ \times & \times & 0 \\ \times & \times & \times \end{bmatrix}, \quad W_2 = \begin{bmatrix} \times & 0 & 0 \\ \times & \times & \times \\ \times & 0 & \times \end{bmatrix}. \quad (10)$$

The associated networks G_{W_1} and G_{W_2} are shown in Fig. 3. Notice that G_{W_1} has more communication links as compared with G_{W_2} (see Remark (R3)).

We consider random numbers for non-zeros in, A, W^1, and C. We have $\rho(A) = 1.3782$, which implies that system is unstable. We determine the block-diagonal gain matrix, K, as in Remark (R4). The system and output noise are considered to be $\mathbf{v}_k \sim \mathbb{N}(0, 0.05^2 I_{n\times n})$ and $\mathbf{r}_k^i \sim \mathbb{N}(0, 0.2^2)$, respectively. In Fig. 5, we compare system error evolution for agents with the Centralized Kalman Filter. We choose all the non-zero parameters randomly and repeat the simulation for 1000 Monte-Carlo trials while recording the sum of squared errors over the $n = 6$ states (for each agent) in each trial. the error is averaged over the Monte-Carlo iterations and then normalized. Clearly, despite the fact that system is unstable, the estimation error at all agents is bounded, even for agent y_C with no system observation. Notice that the performance in terms of estimation error differs for agents.

[1]Notice that the matrix W has to be stochastic.

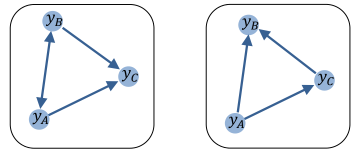

Fig. 3. Two possible communication networks: G_{W_1} on the left and G_{W_2} at on the right;

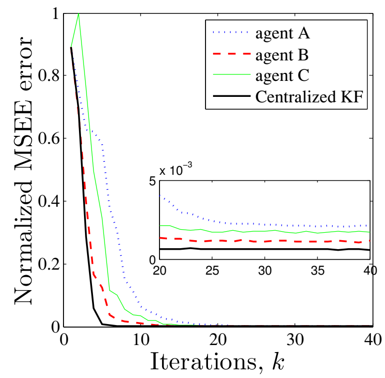

Fig. 4. The performance of the DKE under the network G_{W_2} compared to Centralized Kalman Filter. The MSEE error is normalized.

Although the results here illustrated with a simple academic example, the algorithms are scalable and practically feasible for any large-scale system. The determination of maximal set $\mathcal{L}$ can be done via combinatorial algorithms to find the maximal matching in associated bipartite graph of the system; e.g. *Hopcraft-Karp* algorithm with running time of $\mathcal{O}(n^{2.5})$ [24]. Also, the parent/child SCC classification can be performed using *DFS algorithm* in $\mathcal{O}(n^2)$ [25]. Therefore, the computation effort of is clearly polynomial.

6. CONCLUSION

In this paper, we study the distributed estimator in Eqs. (3)-(4) for S-rank deficient systems; we show that fusion in predictor-space does not result in an observable estimator, and one has to rely on fusion in observation-space. Further, using a generic approach, our results are independent of any particular fusion rule chosen (e.g., Metropolis-Hastings [26]). It is noteworthy that the employed algorithms for agent classification do not have exponential complexity, and thus, the proposed strategies are computationally efficient for large-scale systems.

7. REFERENCES

[1] R. Olfati-Saber, "Distributed Kalman filters with embedded consensus filters," in *44th IEEE Conference on Decision and Control*, Seville, Spain, Dec. 2005, pp. 8179–8184.

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

[3] R. Carli, A. Chiuso, L. Schenato, and S. Zampieri, "Distributed Kalman filtering using consensus strategies," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 4, pp. 622–633, May 2008.

[4] E. J. Msechu, S. D. Roumeliotis, A. Ribeiro, and G. B. Giannakis, "Decentralized quantized Kalman filtering with scalable communication cost," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3727–3741.

[5] U. A. Khan, S. Kar, A. Jadbabaie, and J. M. F. Moura, "On connectivity, observability, and stability in distributed estimation," in *49th IEEE Conference on Decision and Control*, 2010, pp. 6639–6644.

[6] U. A. Khan and A. Jadbabaie, "Coordinated networked estimation strategies using structured systems theory," in *49th IEEE Conference on Decision and Control*, Orlando, FL, Dec. 2011, pp. 2112–2117.

[7] M. Doostmohammadian and U. A. Khan, "Communication strategies to ensure generic networked observability in multiagent systems," in *45th Annual Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2011, pp. 1865–1868.

[8] M. Farina, G. Ferrari-Trecate, and R. Scattolini, "Distributed moving horizon estimation for linear constrained systems," *IEEE Transactions on Automatic Control*, vol. 55, no. 11, pp. 2462–2475, Nov. 2010.

[9] G. Battistelli, L. Chisci, S. Morrocchi, and F. Papi, "An information-theoretic approach to distributed state estimation," in *18th IFAC World Congress*, 2011, pp. 12477–12482.

[10] S. Tu and A. H. Sayed, "Diffusion strategies outperform consensus strategies for distributed estimation over adaptive networks," *IEEE Transactions on Signal Processing,*, vol. 60, no. 12, pp. 6217–6234, Dec. 2012.

[11] T. Boukhboza, F. Hamelin, S. Martinez-Martinez, and D. Sauter, "Structural analysis of the partial state and input observability for structured linear systems: Application to distributed systems," *European Journal of Control*, vol. 15, no. 5, pp. 503–516, Oct. 2009.

[12] A. Wolfgang, J. Akhtman, S. Chen, and L. Hanzo, "Iterative MIMO detection for rank-deficient systems," *IEEE Signal Processing Letters*, vol. 13, no. 11, pp. 699–702, Nov. 2006.

[13] A. Amiri-Simkooei, "Formulation of norm minimization in gauss-markov models," *Journal of Surveying Engineering*, vol. 129, no. 1, pp. 37–43, 2003.

[14] U. A. Khan and M. Doostmohammadian, "A sensor placement and network design paradigm for future smart grids," in *4th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, San Juan, Puerto Rico, Dec. 2011, pp. 137–140.

[15] I. A. Hiskens, "Dynamics of Type-3 wind turbine generator models," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 465–474, Feb. 2012.

[16] C. Commault, J. M. Dion, and D. Trinh, "Observability preservation under sensor failure," *IEEE Transactions on Automatic Control*, vol. 53, no. 6, pp. 1554–1559, Jul. 2008.

[17] J. M. Dion, C. Commault, and J. van der Woude, "Generic properties and control of linear structured systems: A survey," *Automatica*, vol. 39, pp. 1125–1144, Mar. 2003.

[18] K. J. Reinschke, *Multivariable control, a graph theoretic approach*, Berlin: Springer, 1988.

[19] Y. Y. Liu, J. J. Slotine, and A. L. Barabási, "Controllability of complex networks," *Nature*, vol. 473, no. 7346, pp. 167–173, May 2011.

[20] Y. Y. Liu, J. J. Slotine, and A. L. Barabási, "Observability of complex systems," *Proceedings of the National Academy of Sciences*, vol. 110, no. 7, pp. 2460–2465, 2013.

[21] R. Kalman and R. Bucy, "New results in linear filtering and prediction theory," *ASME Journal of Basic Engineering*, vol. 83, pp. 95–108, 1961.

[22] C. Godsil and G. Royle, *Algebraic graph theory*, New York:Springer, 2001.

[23] M. Doostmohammadian and U. Khan, "On the genericity properties in distributed estimation: Topology design and sensor placement," *IEEE Journal of Selected Topics in Signal Processing: Adaptation and Learning over Complex Networks*, vol. 7, no. 2, 2013.

[24] J. Hopcroft and R. Karp, "An $n^{5/2}$ algorithm for maximum matchings in bipartite graphs," *SIAM Journal of Computation*, vol. 2, pp. 225–231, 1973.

[25] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, MIT Press, 2009.

[26] L. Xiao, S. Boyd, and S. Kim, "Distributed average consensus with least-mean-square deviation," *Journal of Parallel and Distributed Computing*, vol. 67, pp. 33–46, 2005.