Automatica 97 (2018) 82–91

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Accuracy analysis for distributed weighted least-squares estimation in finite steps and loopy networks☆

Tianju Sui [a], Damián Edgardo Marelli [b,c,*], Minyue Fu [d,b], Renquan Lu [b]

[a] Department of Control Science and Engineering, Dalian University of Technology, Liaoning, China
[b] School of Automation, Guandong University of Technology, Guangzhou, China
[c] French-Argentinean International Center for Information and Systems Sciences, National Scientific and Technical Research Council, Ocampo y Esmeralda, Rosario 2000, Argentina
[d] School of Electrical Engineering and Computer Science, University of Newcastle, University Drive, Callaghan, NSW 2308, Australia

ARTICLE INFO

Article history:
Received 9 September 2016
Received in revised form 26 December 2017
Accepted 25 June 2018
Available online 17 August 2018

Keywords:
Distributed statistical estimation
Weighted least squares
Convergence rate

ABSTRACT

Distributed parameter estimation for large-scale systems is an active research problem. The goal is to derive a distributed algorithm in which each agent obtains a local estimate of its own subset of the global parameter vector, based on local measurements as well as information received from its neighbors. A recent algorithm has been proposed, which yields the optimal solution (i.e., the one that would be obtained using a centralized method) in finite time, provided the communication network forms an acyclic graph. If instead, the graph is cyclic, the only available alternative algorithm, which is based on iterative matrix inversion, achieving the optimal solution, does so asymptotically. However, it is also known that, in the cyclic case, the algorithm designed for acyclic graphs produces a solution which, although non optimal, is highly accurate. In this paper we do a theoretical study of the accuracy of this algorithm, in communication networks forming cyclic graphs. To this end, we provide bounds for the sub-optimality of the estimation error and the estimation error covariance, for a class of systems whose topological sparsity and signal-to-noise ratio satisfy certain condition. Our results show that, at each node, the accuracy improves exponentially with the so-called loop-free depth. Also, although the algorithm no longer converges in finite time in the case of cyclic graphs, simulation results show that the convergence is significantly faster than that of methods based on iterative matrix inversion. Our results suggest that, depending on the loop-free depth, the studied algorithm may be the preferred option even in applications with cyclic communication graphs.

1. Introduction

With the fast development of sensor networks and wireless communications, the scale of systems is becoming increasingly large. Since centralized estimation requires a fusion center to process all the information from the whole graph, the computation and communication burden increases with the system's size. Thus, the centralized estimation approach is not suitable for large-scale systems, and distributed approaches are needed. The development of distributed estimation has attracted a great deal of attention (Garin & Schenato, 2010; Gupta, Dana, Hespanha, Murray, & Hassibi, 2009; Li & Alregib, 2009; Ribeiro & Giannakis, 2006a, b). It finds applications in industrial monitoring, multi-agent systems, the smart grid, etc.

The distributed estimation problem consists of a network of interconnected nodes, each of which aims to obtain an estimate of certain vector of interest. This is achieved through an iterative procedure in which each node processes its available information, and exchange relevant information with its neighbors, in order to successively compute the required estimate as accurately as possible. The existing distributed estimation problems can be broadly classified into four classes. These classes are: static fully reconstructive, static partially reconstructive, dynamic fully reconstructive and dynamic partially reconstructive. A fully reconstructive system is one in which each node aims to obtain an estimate of the same vector. In contrast, in a partially reconstructive system, each node aims to obtain an estimate of its own partial sub-vector of interest. Also, a static system is one in which prior knowledge of the state at

☆ This work was supported by the China Postdoctoral Science Foundation, China (2017M621134). The material in this paper was presented at the 2015 IEEE 54th Annual Conference on Decision and Control (CDC), December 15–18, 2015. Osaka, Japan. This paper was recommended for publication in revised form by Associate Editor Shreyas Sundaram under the direction of Editor Christos G. Cassandras.

* Corresponding author at: School of Automation, Guandong University of Technology, Guangzhou, China.
E-mail addresses: suitj@mail.dlut.edu.cn (T. Sui), Damian.Marelli@newcastle.edu.au (D.E. Marelli), Minyue.Fu@newcastle.edu.au (M. Fu), rqlu@gdut.edu.cn (R. Lu).

https://doi.org/10.1016/j.automatica.2018.07.016
0005-1098/

a certain time is independent of the knowledge of the same state at previous times. A dynamic system refers to the complementary case. We point out that methods for dynamic estimation can be readily used for static problems, by choosing the dynamic model in a way such that the state stays constant over time.

In the static fully reconstructive problem, the most popular distributed estimation algorithm is consensus (Garin & Schenato, 2010). By running average consensus on the information vector and information matrix of each node, in view of the weighted least squares (WLS) formula, the final estimate of each node converges to the one obtained via WLS (Olfati-Saber, 2005). Although the average consensus algorithm is simple, it has two main disadvantages: First, the communication burden is large, as each node communicates $\frac{n\times(n+3)}{2}$ scalars to its neighbors, where n is the dimension of the estimated vector. Second, the convergence of average consensus requires infinite iterations, and the stopping criterion is still an open problem. To avoid these two disadvantages, many algorithms have been proposed (Ajgl & Šimandl, 2014; Calafiore & Abrate, 2009; Chen, Arambel, & Mehra, 2002; Pasqualetti, Carli, & Bullo, 2012). One of the most important works is the one in Pasqualetti et al. (2012), where using the space structure of measurements and doing kernel projection, each node achieves its minimum norm solution in a finite number of steps.

In the static partially reconstructive problem, since each node considers its own partial state, the consensus algorithm is not applicable. For the case in which the graph induced by the communication network is acyclic (i.e., without loops), an algorithm is proposed in Tai, Lin, Fu, and Sun (2013). In this algorithm, each node obtains a WLS estimate on its own state in a finite number of steps. When the graph is cyclic (i.e., with loops), Marelli and Fu (2015) gave a novel method which, based on Richardson iterations, solves the WLS estimation problem. However, it does so asymptotically, i.e., in infinite iterations. We point out that most estimation algorithms for large-scale systems are partially reconstructive, since the whole state of the system is often of very high dimension.

In the dynamic fully reconstructive problem, the consensus algorithm is also a popular option. In Matei and Baras (2012), one consensus algorithm is run at each sampling time, using the partial estimates obtained at each node, based on their local measurements. Building on this line, a study on the number of consensus iterations required at each sampling time to guarantee the stability of the estimator, under the observability condition, is done in Acikmese, Mandić, and Speyer (2014). Also, the so-called diffusion Kalman filter (Cattivelli & Sayed, 2010) runs consensus on the estimates obtained at each sensor, using local measurements as well as those from neighbors. As opposite to doing consensus on the estimates, the authors of Battistelli and Chisci (2014) found that, by running consensus on the information matrices and vectors, observability is sufficient for the estimation stability.

Concerning the dynamic partially reconstructive problem, information passing and processing methods guaranteeing a stable estimate are proposed in Farina, Ferrari-Trecate, and Scattolini (2010), Khan and Moura (2008), Zhou (2013) and Zhou (2015). Also, the authors of Haber and Verhaegen (2013) study systems with banded dynamic state transition matrices, concluding that the contribution from faraway nodes decreases with the increase of their distance. The authors also propose the moving horizon estimation approach as an approximation to the optimal state estimate.

In this paper we focus on the static partially reconstructive problem. Also, as typically done in static problems, we assume that the vector to be estimated is deterministic. More precisely, we consider the algorithm in Tai et al. (2013), which, as mentioned, yields the optimal solution in finite-time, only when the communication graph is acyclic. For cyclic graphs, this algorithm is not guaranteed to produce the optimal solution. Nevertheless, in many applications, even in the presence of loops, it delivers very good approximations to the optimal solution, in only a very few steps. For those applications, this makes the algorithm a valid alternative to the method in Marelli and Fu (2015) even for cyclic networks. This is because, while the later guarantees the optimal solution, the former one converges much faster. Motivated by this, we study the accuracy of the estimate produced by the algorithm in Tai et al. (2013), under the general setting of a cyclic graph.

For a class of systems whose topological sparsity and signal-to-noise ratio satisfy certain condition, we are able to determine the accuracy of the estimates and their associated estimation error covariances, with respect to those achievable via a centralized WLS method. Our formulas clearly show how accuracy depends on the so-called loop-free depth of each node. More precisely, the estimates and estimation error covariances approach those from the centralized solution, exponentially on the loop-free depth.

The rest of this paper is organized as follows. In Section 2, we give the problem formulation and introduce the distributed WLS algorithm under study. In Section 3, we show how to convert a given graph into other equivalent ones, which are instrumental for analyzing the behavior of the algorithm in cyclic graphs. In Section 4, we introduce our notation, as well as the definition of the *Riemannian Distance* between matrices, together with some of its properties. The accuracy of the information matrices (i.e., the inverses of the error covariances) and state estimates produced by the distributed WLS algorithm are analyzed in Sections 5 and 6, respectively. In Section 7, we provide some simulations to illustrate our results. Finally, concluding remarks are stated in Section 8. Due to space constraints, some complementary mathematical material, including most proofs and some additional lemmas, appears in the extended version (Sui, Marelli, Fu, & Lu, 2018), which is available online.

2. Problem formulation

Consider a system observed by I sensing nodes. Associated to this system, there is a deterministic vector $x^T = \left[x_1^T, x_2^T, \ldots, x_I^T\right] \in \mathbb{R}^n$, with $\sum_{i=1}^{I} n_i = n$, called the global state. For any $i = 1, \ldots, I$, node i aims to estimate the sub-vector $x_i \in \mathbb{R}^{n_i}$. There are also two kinds of measurements. The so-called *self measurements* for node i

$$z_i = C_i x_i + v_i, \tag{1}$$

and the (pair-wise) *joint measurements* between nodes i and j

$$z_{i,j} = C_{i,j} x_i + C_{j,i} x_j + v_{i,j}. \tag{2}$$

In the above, the matrices C_i, $C_{i,j}$ and $C_{j,i}$ are known, and v_i and $v_{i,j}$ are independent measurement noises with known covariances $R_i > 0$ and $R_{i,j} > 0$, respectively. Note that (1) the pair (i, j) is unordered, i.e., $(i, j) = (j, i)$; (2) $z_{i,j} = z_{j,i}$ and $v_{i,j} = v_{j,i}$; (3) It is not necessary for all nodes to have self measurements or all node pairs to have joint measurements. In fact, joint measurements are typically sparse for large graphs.

We assume that node i and node j could communicate if $z_{i,j}$ exists. Furthermore, we call node j a neighbor of node i (i.e., $j \in \mathcal{N}_i$) and node i a neighbor of node j (i.e., $i \in \mathcal{N}_j$) if there is communication between them. In view of this, communication between nodes is always two-ways; and therefore, the associated communication graph (which will be formally introduced later) is always undirected.

The target of distributed WLS estimation is to compute the WLS estimate for each x_i, and its associated estimation error covariance, using a fully distributed algorithm. The algorithm summarized in Algorithm 1, achieves this goal. In this algorithm, at iteration N, node i computes a local estimate $\hat{x}_i(N)$ of its sub-vector of interest,

Algorithm 1 Distributed WLS algorithm.

1) **Initialization:** At time $k = 0$, node i defines:

$$\alpha_{j\to i}(0) = 0, \quad Q_{j\to i}(0) = 0. \tag{3}$$

2) **Main loop:** At time $N = 1, 2, \cdots$, do:
2.1) Each node i computes

$$\alpha_i(N) = C_i^T R_i^{-1} z_i + \sum_{j\in\mathcal{N}_i} \alpha_{j\to i}(N-1),$$
$$Q_i(N) = C_i^T R_i^{-1} C_i + \sum_{j\in\mathcal{N}_i} Q_{j\to i}(N-1), \tag{4}$$

and

$$\hat{x}_i(N) = Q_i^{-1}(N)\alpha_i(N), \quad \Sigma_i(N) = Q_i^{-1}(N). \tag{5}$$

2.2) Each node i sends to each connected node j with $j \in \mathcal{N}_i$:

$$\alpha_{i\to j}(N) = C_{j,i}^T R_{i\to j}^{-1}(N) z_{i\to j}(N),$$
$$Q_{i\to j}(N) = C_{j,i}^T R_{i\to j}^{-1}(N) C_{j,i}, \tag{6}$$

where

$$z_{i\to j}(N) = z_{i,j} - C_{i,j}\left(Q_i(N) - Q_{j\to i}(N-1)\right)^{-1} \cdot (\alpha_i(N) - \alpha_{j\to i}(N-1)), \tag{7}$$
$$R_{i\to j}(N) = R_{i,j} + C_{i,j}(Q_i(N) - Q_{j\to i}(N-1))^{-1} C_{i,j}^T. \tag{8}$$

and its associated covariance $\Sigma_i(N)$, using its local information vector $\alpha_i(N)$ and information matrix $Q_i(N)$. Then, for each neighbor $j \in \mathcal{N}_i$, it removes from $\alpha_i(N)$ and $Q_i(N)$ the information vector $\alpha_{j\to i}(N-1)$ and matrix $Q_{j\to i}(N-1)$, respectively, which it received at the previous iteration from neighbor j, to built the information vector $\alpha_{i\to j}(N)$ and matrix $Q_{i\to j}(N)$, that it sends to the same neighbor at the current iteration.

Algorithm 1 requires Assumption 2, which is given below. This assumption implies that each node is able to obtain an (possibly coarse) estimate of its sub-vector of interest, using only its self measurements. Notice that, if this assumption is not met, we have, at time $N = 1$, and node i, that $Q_i(1) - Q_{j\to i}(0) = C_i^T R_i^{-1} C_i$. Hence, $Q_i(1) - Q_{j\to i}(0)$ cannot be inverted in (7) and (8) .

Before stating Assumption 2 we introduce some required notation.

Notation 1. *The superscript ^{T} denotes vector or matrix transposition. For a matrix A, $\|A\|$ denotes the induced operator norm, i.e., the maximum singular value of A. Also, $A > 0$ ($A \geq 0$) means that A is positive definite (semi-definite), i.e., $x^T Ax > 0$ ($x^T Ax \geq 0$), for all $x \neq 0$. For a second matrix B, $A > B$ ($A \geq B$) means that $A - B > 0$ ($A - B \geq 0$).*

Assumption 2. For every $i = 1, 2, \ldots, I$, we have

$$C_i^T R_i^{-1} C_i > 0.$$

It is known that Algorithm 1 converges to the correct estimates in a finite number of iterations, when its associated communication graph is acyclic (Tai et al., 2013). In fact, the required number of iterations equals the diameter of the graph, i.e., the maximum number of edges connecting one node to another over the whole graphs. The fundamental challenge in our study is to understand how the algorithm performs for cyclic graphs. As mentioned in Section 1, the goal of this paper is to quantify the accuracy of the estimate when the graph is cyclic, i.e., quantify the difference between the distributed estimate and the centralized one. We split our accuracy analysis in that of the information matrix (Section 5) and that of the state estimate (Section 6). In the rest of paper, without loss of generality, we concentrate our study on the accuracy of an arbitrary node, which is labeled as node 1.

3. Graph representations

In the network described above, nodes have only self and pairwise joint measurements. This permits using a simple connectivity (undirected) graph, called *canonical graph*, to describe the nodes and their measurements. This is explained in Section 3.1. A drawback of this representation for our intended analysis is that this graph is cyclic in general. In Section 3.2 we describe how to convert this cyclic graph into an acyclic one, with an infinite number of nodes. It turns out that this is an equivalent graph, as far as distributed estimation is concerned. Then, in Section 3.3, we explain how to further convert the acyclic graph into another equivalent one, whose topology is that of a single line. For a more detailed presentation of equivalent graph transformations, the reader is referred to Tatikonda and Jordan (2002), Tatikonda (2003) and Weiss (2000). We point out that all the above graphs are undirected.

3.1. Canonical graph representation

The canonical graph $\mathcal{G}$ has a node associated with each sensing node $i = 1, \ldots, I$. Also, nodes i and j are connected by an undirected edge if they can communicate with each other, i.e., $j \in \mathcal{N}_i$.

3.2. Acyclic graph representation

We start the section with the following definition.

Definition 3. A rooted tree graph is an acyclic connected graph, in which a node is assigned as its root. For a node i different from the root one, we let $p(i)$ denote its *parent* (i.e., the next node when moving towards the root) and $\mathcal{S}_i$ denote the set of its *children* (i.e., all the nodes j with $i = p(j)$). Also, node i is called a leaf if $\mathcal{S}_i$ is empty.

Given a cyclic canonical graph $\mathcal{G}$, we can convert it into an acyclic one $\mathcal{A}$, having a rooted tree topology, with any arbitrary node as its root one. Since, as pointed out before, we concentrate our study on the accuracy at node 1, we choose this node as the root one. This graph enjoys the property that, if Algorithm 1 is applied to both graphs, it will produce at node 1 and iteration N, the same result.

Graph $\mathcal{A}$ has an infinite number of nodes. Each of its nodes is associated to a node in $\mathcal{G}$. With some abuse of notation, we use $\mathcal{G}(n)$ to denote the node in $\mathcal{G}$ associated to node n in $\mathcal{A}$, and $\mathcal{A}(i)$ to denote the set of nodes in $\mathcal{A}$ associated to node i in $\mathcal{G}$. Graph $\mathcal{A}$ is constructed as the limit of the following iterative procedure. We start by defining $\mathcal{A}_0$ as the empty graph and $\mathcal{A}_1$ as the graph having no edges, and having a single node, which is associated to node 1 of $\mathcal{G}$. Then, at each step $N \geq 2$, we do the following steps:

(1) Find all leaf nodes l of the tree $\mathcal{A}_{N-1}$.
(2) Find all neighbors j of $\mathcal{G}(l)$ in $\mathcal{G}$, excluding all nodes in $\mathcal{A}(p(l))$ (i.e., associated to the parent of l in $\mathcal{A}$).
(3) For each l and j:
   (a) Add a node n to the tree.
   (b) Add the undirected edge from l to n to the tree.
   (c) Associate n in $\mathcal{A}$ to j in $\mathcal{G}$.
(4) Define $\mathcal{A}_N$ as the resulting graph.

Our next step is to associate a system of measurement equations to the acyclic graph $\mathcal{A}$, in a way similar to the way in which the system of Eqs. (1)–(2) is associated with $\mathcal{G}$. These equations need to satisfy two conditions. First, their canonical graph should be $\mathcal{A}$. Second, the aforementioned equivalence at node 1 should be preserved. As explained in Ihler, Iii, and Willsky (2005), Tatikonda (2003), Tatikonda and Jordan (2002) and Weiss (2000), both conditions are satisfied if

$$\bar{z}_i = \bar{C}_i \bar{x}_i + \bar{v}_i, \tag{9}$$

$$\bar{z}_{i,j} = \bar{C}_{i,j} \bar{x}_i + \bar{C}_{j,i} \bar{x}_j + \bar{v}_{i,j} \tag{10}$$

for all $i \in \mathcal{A}$ and $j \in \mathcal{S}_i$, with $\bar{v}_i \sim \mathcal{N}\left(0, \bar{R}_i\right)$ and $\bar{v}_{i,j} \sim \mathcal{N}\left(0, \bar{R}_{i,j}\right)$. The values of the quantities in (9)–(10) are given by those corresponding to the nodes in $\mathcal{G}$ which are associated to nodes i and j in $\mathcal{A}$, i.e.,

$$\bar{z}_i = z_{\mathcal{G}(i)}, \quad \bar{v}_i = v_{\mathcal{G}(i)}, \quad \bar{C}_i = C_{\mathcal{G}(i)},$$
$$\bar{R}_i = R_{\mathcal{G}(i)}, \quad \bar{z}_{i,j} = z_{\mathcal{G}(i),\mathcal{G}(j)}, \quad \bar{v}_{i,j} = v_{\mathcal{G}(i),\mathcal{G}(j)},$$
$$\bar{C}_{i,j} = C_{\mathcal{G}(i),\mathcal{G}(j)}, \quad \bar{R}_{i,j} = R_{\mathcal{G}(i),\mathcal{G}(j)}.$$

Also, all noises $\bar{v}_i$ and $\bar{v}_{i,j}$, $i = 1, \ldots, I$, $j \in \mathcal{N}_i$, are pairwise uncorrelated.

Remark 4. Recall that, in an acyclic graph, Algorithm 1 produces, at each node, the same estimate that would be obtained using the centralized WLS method. Since the graph $\mathcal{A}_N$ is acyclic, the outcomes of both methods will be the same on $\mathcal{A}_N$. Moreover, its measurements (9)–(10), are designed so that, at node 1 and iteration N, this outcome equals that resulting from applying Algorithm 1 to graph $\mathcal{G}$. Hence, $\mathcal{G}$ and $\mathcal{A}_N$ are equivalent graphs only from the point of view of Algorithm 1 (at node 1 and iteration N), but not from that of centralized WLS.

3.3. Representation as a line graph

Let $\mathcal{A}_N$ be the N-layer acyclic graph with root node 1 and measurement equations (9)–(10), as described above. We now describe how to convert $\mathcal{A}_N$ into a line graph $\mathcal{L}_N$ such that the aforementioned equivalence is still preserved.

Indeed, $\mathcal{L}_N$ is formed by simply grouping all the nodes in $\mathcal{A}_N$, which are exactly $n-1$ hops away from node 1, into a super node $\mathcal{T}_n$, for all $n = 1, 2, \ldots, N$. In particular, $\mathcal{T}_1$ is just node 1.

Again, we need to associate a system of measurement equations to $\mathcal{L}_N$ satisfying the conditions described in Section 3.2. This is done by grouping all the measurement equations for each super node $\mathcal{T}_n$, as detailed below.

Denote the size of any finite set S by $|S|$ and its elements by $S(1), S(2), \ldots, S(|S|)$. For each $n \in \mathbb{N}$, the state of $\mathcal{T}_n$ is given by

$$\tilde{x}_n = \left[\bar{x}^T_{\mathcal{T}_n(1)}, \quad \bar{x}^T_{\mathcal{T}_n(2)}, \quad \ldots, \quad \bar{x}^T_{\mathcal{T}_n(|\mathcal{T}_n|)}\right]^T,$$

and its measurement equations are given by

$$\tilde{z}_n = \tilde{C}_n \tilde{x}_n + \tilde{v}_n, \tag{11}$$

$$\tilde{z}_{n,n+1} = \tilde{C}_{n,n+1} \tilde{x}_n + \tilde{C}_{n+1,n} \tilde{x}_{n+1} + \tilde{v}_{n,n+1}. \tag{12}$$

That is, $\tilde{z}_n$ consists of all the measurements $\bar{z}_i$ with $i \in \mathcal{T}_n$, and $\tilde{z}_{n,n+1}$ consists of all the measurements $\bar{z}_{i,j}$ with $i \in \mathcal{T}_n$ and $j \in \mathcal{T}_{n+1}$. Note that $\tilde{v}_n \sim \mathcal{N}(0, \tilde{R}_n)$ and $\tilde{v}_{n,n+1} \sim \mathcal{N}(0, \tilde{R}_{n,n+1})$. The matrices $\tilde{C}_n$, $\tilde{R}_n$, $\tilde{C}_{n,n+1}$ and $\tilde{R}_{n,n+1}$ are naturally related to $\bar{C}_i$, $\bar{R}_i$, $\bar{C}_{i,j}$ and $\bar{R}_{i,j}$ through the above construction. More precisely,

$$\tilde{z}_n = \left[\bar{z}^T_{\mathcal{T}_n(1)}, \quad \bar{z}^T_{\mathcal{T}_n(2)}, \quad \ldots, \quad \bar{z}^T_{\mathcal{T}_n(|\mathcal{T}_n|)}\right]^T,$$
$$\tilde{C}_n = \operatorname{diag}\{\bar{C}_{\mathcal{T}_n(1)}, \bar{C}_{\mathcal{T}_n(2)}, \ldots, \bar{C}_{\mathcal{T}_n(|\mathcal{T}_n|)}\},$$
$$\tilde{R}_n = \operatorname{diag}\{\bar{R}_{\mathcal{T}_n(1)}, \bar{R}_{\mathcal{T}_n(2)}, \ldots, \bar{R}_{\mathcal{T}_n(|\mathcal{T}_n|)}\}.$$

Similarly,

$$\tilde{z}_{n,n+1} = \left[\grave{z}^T_{\mathcal{T}_n(1)}, \quad \grave{z}^T_{\mathcal{T}_n(2)}, \quad \ldots, \quad \grave{z}^T_{\mathcal{T}_n(|\mathcal{T}_n|)}\right]^T,$$
$$\tilde{C}_{n,n+1} = \operatorname{diag}\{\grave{C}_{\mathcal{T}_n(1)}, \grave{C}_{\mathcal{T}_n(2)}, \ldots, \grave{C}_{\mathcal{T}_n(|\mathcal{T}_n|)}\},$$
$$\tilde{C}_{n+1,n} = \operatorname{diag}\{\acute{C}_{\mathcal{T}_{n+1}(1)}, \acute{C}_{\mathcal{T}_{n+1}(2)}, \ldots, \acute{C}_{\mathcal{T}_{n+1}(|\mathcal{T}_{n+1}|)}\},$$
$$\tilde{R}_{n,n+1} = \operatorname{diag}\{\grave{R}_{\mathcal{T}_n(1)}, \grave{R}_{\mathcal{T}_n(2)}, \ldots, \grave{R}_{\mathcal{T}_n(|\mathcal{T}_n|)}\},$$

with

$$\grave{z}_i = \left[\bar{z}^T_{i,\mathcal{S}_i(1)}, \quad \bar{z}^T_{i,\mathcal{S}_i(2)}, \quad \ldots, \quad \bar{z}^T_{i,\mathcal{S}_i(|\mathcal{S}_i|)}\right]^T,$$
$$\grave{C}_i = \left[\bar{C}^T_{i,\mathcal{S}_i(1)}, \quad \bar{C}^T_{i,\mathcal{S}_i(2)}, \quad \ldots, \quad \bar{C}^T_{i,\mathcal{S}_i(|\mathcal{S}_i|)}\right]^T,$$
$$\acute{C}_i = \operatorname{diag}\{\bar{C}_{\mathcal{S}_i(1),i}, \bar{C}_{\mathcal{S}_i(2),i}, \ldots, \bar{C}_{\mathcal{S}_i(|\mathcal{S}_i|),i}\},$$
$$\grave{R}_i = \operatorname{diag}\{\bar{R}_{i,\mathcal{S}_i(1)}, \bar{R}_{i,\mathcal{S}_i(2)}, \ldots, \bar{R}_{i,\mathcal{S}_i(|\mathcal{S}_i|)}\}.$$

Remark 5. Note that the statement in Remark 4 also holds for $\mathcal{L}_N$. More precisely, the centralized WLS estimate at node 1 in $\mathcal{L}_N$ equals to that of Algorithm 1, when applied to $\mathcal{G}$, at the same nodes and iteration N.

In view of the above analysis, the problem of studying the dynamics of Algorithm 1 becomes the problem of studying the centralized WLS estimate at node 1 for the graph $\mathcal{L}_N$, as $N \to \infty$.

4. Preliminaries

In our analysis below, we will make use of the so-called *Riemannian distance* between matrices (Bougerol, 1993).

Definition 6. For $n \times n$ matrices $P, Q > 0$, their Riemannian distance is defined by

$$\delta(P, Q) = \sqrt{\sum_{k=1}^{n} \log^2 \sigma_k\left(PQ^{-1}\right)},$$

where $\sigma_1(X) \geq \cdots \geq \sigma_n(X)$ denote the singular values of matrix X.

The following proposition states a number of properties of the Riemannian distance. Its proof appears in Sui et al. (2018, Appendix).

Proposition 7. *For any $n \times n$ positive definite matrices P and Q, the following results hold:*

(1) $\delta(P, P) = 0$.
(2) $\delta\left(P^{-1}, Q^{-1}\right) = \delta(Q, P) = \delta(P, Q)$.
(3) If B has full row rank, $\delta\left(BPB^T, BQB^T\right) \leq \delta(P, Q)$, and the equality holds if B is invertible.
(4) If $P \geq Q$ and $W \geq 0$, then $\delta(P + W, Q) \geq \delta(P, Q)$.
(5) For any $m \times m$ matrix $W > 0$ and $m \times n$ matrix B, we have

$$\delta(W + BP^{-1}B^T, W + BQ^{-1}B^T) \leq \frac{\alpha}{\alpha + \beta}\delta(P, Q),$$

where $\alpha = \max\{\|BP^{-1}B^T\|, \|BQ^{-1}B^T\|\}$ and $\beta = \sigma_{\min}(W)$, with $\sigma_{\min}(W)$ denoting the smallest singular value of W.
(6) If $P > Q$, then $\|P - Q\| \leq \left(e^{\delta(P,Q)} - 1\right)\|Q\|$.

We now introduce some notation that will be used in the rest of the paper.

Notation 8. *For a graph $\mathcal{C}$, we use $\hat{x}_i(\mathcal{C})$, $Q_i(\mathcal{C})$ and $\alpha_i(\mathcal{C})$ to denote the final (after convergence) state estimate, information matrix and information vector, respectively, obtained by running Algorithm 1 on $\mathcal{C}$.*

Notation 9. *Let $Q_{(i,j),0} \geq 0$, for each $i = 1, \ldots, I$, $j \in \mathcal{N}_i$, and define the set $\mathbb{Q} = \left\{Q_{(i,j),0} : i = 1, 2, \ldots, I, j \in \mathcal{N}_i\right\}$. Suppose that we run Algorithm 1 on the network $\mathcal{G}$, but replacing the initialization (3) by $Q_{j\to i}(0) = Q_{(i,j),0}$, for all $i = 1, \ldots, I$, $j \in \mathcal{N}_i$. We use $\hat{x}_1(N, \mathbb{Q})$ and $Q_1(N, \mathbb{Q})$ to denote the estimate and information matrix, respectively, yielded by such algorithm at node 1 and step N.*

Notation 10. *If $z_{i,j}$ exists, the pair (i, j) is called an (undirected)* edge. *Notice that since edges are undirected, the pairs (i, j) and (j, i) denote the same edge. A* path *is a concatenation of contiguous edges, and its* length *is the number of edges forming it. A* cycle *is a path with no repetitions of vertices and edges, except for the necessary repetition of the starting and ending vertices. For each $i, j \in \{1, \ldots, I\}$, the* distance *between nodes i and j is defined as the minimum length of a path joining these two nodes. Let $\mathcal{N}_1(l)$ denote the subgraph of $\mathcal{G}$ formed by nodes whose distance from node 1 is less than or equal to l. The loop-free depth l_1 of node 1 is the largest integer such that $\mathcal{N}_1(l_1)$ is acyclic (i.e., without* cycles*).*

We also introduce the following constants

$$\bar{u} := \max_i |\mathcal{N}_i| - 1, \quad \bar{n} := \max_i \dim x_i,$$
$$\bar{m} := \max\{\max_i \dim z_i, \max_{i,j} \dim z_{i,j}\}.$$

5. Accuracy analysis for the information matrix

In this section we derive a bound for the difference between the information matrix yielded by Algorithm 1 and that obtained using the centralized WLS method. For a class of systems, we provide a lower bound for the time after which the difference falls within this bound. Moreover, this bound decreases exponentially with the increase of the loop free-depth. Our main result is given in Section 5.1. Its proof appears in Section 5.2.

5.1. Main result

The main result on the accuracy of the information matrix is given below.

Theorem 11. *Let $\mathrm{Cov}_1^{\mathrm{WLS}}$ be the estimation error covariance obtained at node 1 when using centralized WLS. If $\rho < 1$, then there exists a constant ϖ (only dependent on the system parameters $\bar{u}$, $\bar{n}$, $C_{i,j}$, C_i, $R_{i,j}$ and R_i) such that, for any $N \geq l_1 + 1$,*

$$\|\mathrm{Cov}_1^{\mathrm{WLS}} - Q_1^{-1}(N)\| \leq \varpi \rho^{l_1},$$

where

$$\rho = \lambda\sqrt{\bar{u}}, \quad \lambda = \frac{\alpha_1}{\alpha_1 + \beta_1}\frac{\alpha_2}{\alpha_2 + \beta_2},$$
$$\alpha_1 = \bar{u} \max_{i,j} \|C_{i,j}^T R_{i,j}^{-1} C_{i,j}\|, \quad \beta_1 = \min_i \sigma_{\min}(C_i^T R_i^{-1} C_i),$$
$$\alpha_2 = \max_{i,j} \|C_{i,j}(C_i^T R_i^{-1} C_i)^{-1} C_{i,j}^T\|, \beta_2 = \min_{i,j} \sigma_{\min}(R_{i,j}).$$

Remark 12. Theorem 11 states that, if the graph/system satisfies $\rho < 1$, the inverse of the information matrix $Q_1^{-1}(N)$ yielded by Algorithm 1 at node 1, exponentially approaches the estimation error covariance of centralized WLS, as its loop-free depth l_1 increases.

Remark 13. Since $\rho = \lambda\sqrt{\bar{u}}$, the result is mainly given for a class of graphs with sparse connections (small $\bar{u}$) and where the ratio between the signal-to-noise-ratio (SNR) of local measurements and the SNR of each joint measurement is small (small λ). Notice that the later condition is relatively mild, since each joint measurement typically has a lower SNR than the local one.

5.2. Proof of Theorem 11

Recall that, in view of the graph conversions described in Section 3, we have $Q_1(N) = Q_1(\mathcal{A}_N) = Q_1(\mathcal{L}_N)$. The proof of Theorem 11 uses this fact. It also requires the following lemmas, whose proofs appear in the (Sui et al. (2018) Appendix).

Lemma 14. *Let $\mathbb{Q}_1$ and $\mathbb{Q}_2$ be initial sets both satisfying $0 \leq Q_{(i,j),0}^c \leq C_{i,j}^T R_{i,j}^{-1} C_{i,j}$, for all $Q_{(i,j),0}^c \in \mathbb{Q}_c$ and $c \in \{1, 2\}$. Then, in the notation of Theorem 11, for all $N \in \mathbb{N}$,*

$$\delta\left(Q_1(N, \mathbb{Q}_1) - Q_1(N, \mathbb{Q}_2)\right) \leq \rho^{N-1}\bar{\delta},$$

with

$$\bar{\delta} = \sqrt{(\bar{u}+1)\bar{n}} \times \max_i \log\|I + (\sum_{j\in\mathcal{N}_i} C_{i,j}^T R_{i,j}^{-1} C_{i,j})(C_i^T R_i^{-1} C_i)^{-1}\|.$$

Notation 15. *Let*

$$\mathbb{Q}^{\mathrm{M}} = \left\{Q_{(i,j),0}^{\mathrm{M}} : i = 1, \ldots, I \text{ and } j \in \mathcal{N}_i\right\},$$

$$\mathbb{Q}^0 = \left\{Q_{(i,j),0}^0 : i = 1, \ldots, I \text{ and } j \in \mathcal{N}_i\right\},$$

with

$$Q_{(i,j),0}^{\mathrm{M}} = C_{i,j}^T R_{i,j}^{-1} C_{i,j}, \quad Q_{(i,j),0}^0 = 0.$$

In particular, notice that $\mathbb{Q}^0$ is the initialization used in Algorithm 1, i.e., $\hat{x}_1(N) = \hat{x}_1(N, \mathbb{Q}^0)$ and $Q_1(N) = Q_1(N, \mathbb{Q}^0)$.

Lemma 16. *Recall the definition of loop-free depth l_1 from Notation 10. For any $N \geq l_1 + 1$, we have*

$$\left\|Q_1^{-1}(N) - \mathrm{Cov}_1^{\mathrm{WLS}}\right\| \leq \left\|Q_1^{-1}(l_1 + 1, \mathbb{Q}^{\mathrm{M}}) - Q_1^{-1}(l_1 + 1, \mathbb{Q}^0)\right\|. \tag{13}$$

The proof of Theorem 11 uses the above property to provide an upper bound for the difference between $Q_1^{-1}(N)$ and $\mathrm{Cov}_1^{\mathrm{WLS}}$.

Proof of Theorem 11. Since both $\mathbb{Q}^{\mathrm{M}}$ and $\mathbb{Q}^0$ satisfy the condition in Lemma 14, it follows that

$$\delta(Q_1^{-1}(l_1 + 1, \mathbb{Q}^0), Q_1^{-1}(l_1 + 1, \mathbb{Q}^{\mathrm{M}})) \leq \rho^{l_1}\bar{\delta}.$$

From Proposition 7,

$$\begin{aligned}&\|Q_1^{-1}(l_1 + 1, \mathbb{Q}^0), Q_1^{-1}(l_1 + 1, \mathbb{Q}^{\mathrm{M}})\| \\ &\leq (e^{\rho^{l_1}\bar{\delta}} - 1)\|Q_1^{-1}(l_1 + 1)\| \leq (e^{\rho^{l_1}\bar{\delta}} - 1)\|(C_1^T R_1^{-1} C_1)^{-1}\| \\ &\leq \beta_1^{-1}(e^{\bar{\delta}} - 1)\rho^{l_1},\end{aligned}$$

where the last inequality follows from Sui et al. (2018, Lemma 30). Since the quantity $\beta_1^{-1}(e^{\bar{\delta}} - 1)$ depends only on $\bar{u}$, $\bar{n}$ and the system parameters $C_{i,j}$, C_i, $R_{i,j}$, R_i for some i and j, the result then follows from (13).

6. Accuracy analysis for the state estimate

In this section, we derive a bound for the difference between the estimate yielded by Algorithm 1 and that obtained using the centralized WLS method. For a class of systems, we provide a lower bound for the time after which the difference falls within this bound. Moreover, this bound decreases exponentially with the increase of the loop-free depth. Our main result is given in Section 6.1 and its proof appears in Section 6.2.

6.1. Main result

The main result on the accuracy of the estimate is given below.

Theorem 17. *Let $\hat{x}_1^{\mathrm{WLS}}$ be the estimate obtained at node 1 when using centralized WLS. If $\kappa < 1$, then there exists a constant ϖ (only dependent on the system parameters $\bar{m}$, $\bar{u}$, $\bar{n}$, $C_{i,j}$, C_i, $R_{i,j}$ and R_i as well as on the measurements $z_{i,j}$, z_i) such that, for all $N \geq l_1 + 1$,*

$$\|\hat{x}_1(N) - \hat{x}_1^{\mathrm{WLS}}\| \leq \varpi \kappa^{l_1+1}$$

where with

$$\kappa = \max\{\bar{u}\sqrt{\omega}, \sqrt{\bar{u}}\iota^{1/\zeta}\}, \quad \omega = \frac{a_1}{a_1+b_1}\frac{a_2}{a_2+b_2},$$

$$a_1 = \underline{r}^{-1}\bar{u}\max_{i,j}\left\|C_{i,j}\right\|^2, \quad a_2 = \max_{i,j}\|C_{i,j}\|^2\frac{\bar{u}\bar{r}}{\underline{\varepsilon}^2},$$

$$b_1 = \bar{r}^{-1}\underline{\varepsilon}^2, \quad b_2 = \underline{r}, \quad \iota = \frac{\sqrt{\bar{q}}-\sqrt{\underline{q}}}{\sqrt{\bar{q}}+\sqrt{\underline{q}}},$$

$$\bar{q} = \bar{\varepsilon}^2\underline{r}^{-1}, \quad \underline{q} = \underline{\varepsilon}^2\bar{r}^{-1}, \quad \zeta = 2 + \log_{\frac{1}{\sqrt{\omega}}}(\bar{q}/\underline{q}).$$

$$\bar{r} = \max_i\{\|R_i\|, \|R_{i,j}\|\}, \underline{r} = \min_i\{\sigma_{\min}(R_i), \sigma_{\min}(R_{i,j})\},$$

$$\bar{\varepsilon} = \max_{i,j}\sqrt{\|C_i\|^2 + 4\bar{u}\left\|C_{i,j}\right\|^2}, \quad \underline{\varepsilon} = \min_i \sigma_{\min}(C_i).$$

Remark 18. From the definition of κ, the observation made in Remark 13 also applies to Theorem 17. However, since $\kappa \geq \rho$, the condition required for Theorem 17 is stronger than that for Theorem 11.

6.2. Proof of Theorem 17

We split the proof of Theorem 17 into three parts. In Section 6.2.1 we derive a bound for the increment $\hat{x}_1(N+1)-\hat{x}_1(N)$ in a graph with line topology. In Section 6.2.2 we generalize this result for an arbitrary graph. Finally, in Section 6.2.3 we use this result to bound the difference between the state estimate yielded by Algorithm 1 and that obtained using the centralized WLS method.

6.2.1. Bound of the increment in a line graph

Consider a line graph $\mathcal{L}_N$ with measurement equations given by (11)–(12). Let

$$y_i = \left[\tilde{z}_i^T, \tilde{z}_{i,i+1}^T\right]^T, \quad w_i = \left[\tilde{v}_i^T, \tilde{v}_{i,i+1}^T\right]^T,$$

$$A_{i,i} = \begin{bmatrix}\tilde{C}_i \\ \tilde{C}_{i,i+1}\end{bmatrix}, \ A_{i,i+1} = \begin{bmatrix}0 \\ \tilde{C}_{i+1,i}\end{bmatrix}, \ S_i = \begin{bmatrix}\tilde{R}_i & 0 \\ 0 & \tilde{R}_{i,i+1}\end{bmatrix}$$

for any $i = 1, 2, \ldots, N-1$, and

$$y_N = \tilde{z}_N, \quad w_i = \tilde{v}_N, \quad A_{N,N} = \tilde{C}_N, \quad S_N = \tilde{R}_N.$$

Then, (11) and (12) become

$$y_i = A_{i,i}\tilde{x}_i + A_{i,i+1}\tilde{x}_{i+1} + w_i,$$

with $w_i \sim \mathcal{N}(0, S_i)$. We also define

$$\mathbf{x}_N = \left[\tilde{x}_1^T, \ldots, \tilde{x}_N^T\right]^T, \ \mathbf{y}_N = \left[y_1^T, \ldots, y_N^T\right]^T,$$
$$\mathbf{w}_N = \left[w_1^T, \ldots, w_N^T\right]^T,$$
$$[\mathbf{A}_N]_{n,m} = \begin{cases}A_{n,m}, & 0 \leq m-n \leq 1, \\ 0, & \text{otherwise.}\end{cases}$$

We then have

$$\mathbf{y}_N = \mathbf{A}_N\mathbf{x}_N + \mathbf{w}_N,$$

with $\mathbf{w}_N \sim \mathcal{N}(0, \mathbf{S}_N)$ and $\mathbf{S}_N = \mathrm{diag}\{S_1, \ldots, S_N\}$.

The WLS estimate $\hat{\mathbf{x}}_N$ of $\mathbf{x}_N$ is given by

$$\hat{\mathbf{x}}_N = \mathbf{Q}_N^{-1}\mathbf{q}_N,$$

where $\mathbf{q}_N = \mathbf{A}_N^T\mathbf{S}_N^{-1}\mathbf{y}_N = \left[q_1^T, \ldots, q_N^T\right]^T$ with

$$q_i = \begin{cases}A_{i,i}^T S_i^{-1} y_i, & i = 1, \\ A_{i,i}^T S_i^{-1} y_i + A_{i-1,i}^T S_{i-1}^{-1} y_{i-1}, & i > 1,\end{cases} \tag{14}$$

and the (i,j)-th entry $Q_{i,j}$ of $\mathbf{Q}_N = \mathbf{A}_N^T\mathbf{S}_N^{-1}\mathbf{A}_N$ given by

$$\begin{aligned} Q_{i,i} &= \begin{cases}A_{i,i}^T S_i^{-1} A_{i,i}, & i = 1, \\ A_{i,i}^T S_i^{-1} A_{i,i} + A_{i-1,i}^T S_{i-1}^{-1} A_{i-1,i}, & i > 1,\end{cases} \\ Q_{i,i+1} &= A_{i,i}^T S_i^{-1} A_{i,i+1}, \ Q_{i+1,i} = Q_{i,i+1}^T, \\ Q_{i,j} &= 0, \ |i-j| \geq 2. \end{aligned} \tag{15}$$

Let $\Sigma_N = \mathbf{Q}_N^{-1}$ and $[\Sigma_N]_{i,j}$ be its (i,j)-th block. From the inverse formula for band matrices given in Theorem 3.1 of Meurant (1992), it follows that the first block row of Σ_N is given by

$$[\Sigma_N]_{1,j} = \left(\prod_{k=1}^{j-1}\Delta_k^{-1}Q_{k,k+1}\right)\Phi_j^{-1}(N) \tag{16}$$

with

$$\Phi_j(N) = \Gamma_j(N) - Q_{j,j-1}\Delta_{j-1}^{-1}Q_{j-1,j}, \tag{17}$$

$$\Delta_k = \begin{cases}Q_{kk}, & k = 1, \\ Q_{kk} - Q_{k,k-1}\Delta_{k-1}^{-1}Q_{k-1,k}, & k > 1,\end{cases}$$

$$\Gamma_k(N) = \begin{cases}Q_{kk}, & k = N, \\ Q_{kk} - Q_{k,k+1}\Gamma_{k+1}^{-1}(N)Q_{k+1,k}, & k < N,\end{cases}$$

for any $j = 1, 2, \ldots, N$. Then, the first entry $\left[\hat{\mathbf{x}}_N\right]_1$ of $\hat{\mathbf{x}}_N$ is given by

$$\left[\hat{\mathbf{x}}_N\right]_1 = \sum_{j=1}^{N}[\Sigma_N]_{1,j}q_j. \tag{18}$$

Recall that $\hat{x}_1(\mathcal{L}_N) = \left[\hat{\mathbf{x}}_N\right]_1$, it then follows from (18) that

$$\begin{aligned}\left\|\hat{x}_1(\mathcal{L}_{N+1}) - \hat{x}_1(\mathcal{L}_N)\right\| \leq &\sum_{j=1}^{N}\left\|[\Sigma_{N+1}]_{1,j} - [\Sigma_N]_{1,j}\right\|\|q_j\| \\ &+ \left\|[\Sigma_{N+1}]_{1,N+1}\right\|\|q_{N+1}\|.\end{aligned} \tag{19}$$

The main result of this section is given in Lemma 26. It bounds the decay rate in (19). Its proof requires a number of lemmas, which are stated below.

We start by stating bounds for certain quantities, namely, $\mathbf{A}_N$, $\mathbf{Q}_N$, Δ_k, $\Gamma_k(N)$, $\Phi_k(N)$. This is done in Lemmas 19–21. Some of the proofs are given in Sui et al. (2018, Appendix).

Lemma 19. *For any $N \in \mathbb{N}$,*

$$\underline{\tilde{\varepsilon}}I \leq \mathbf{A}_N \leq \bar{\tilde{\varepsilon}}I,$$

with

$$\bar{\tilde{\varepsilon}} = \max_i\Big(\left\|\tilde{C}_i\right\|^2 + 2\max\{\|\tilde{C}_{i-1,i}\|^2, \|\tilde{C}_{i,i-1}\|^2\} + 2\max\{\|\tilde{C}_{i,i+1}\|^2, \|\tilde{C}_{i+1,i}\|^2\}\Big)^{1/2},$$

$$\underline{\tilde{\varepsilon}} = \min_i \sigma_{\min}(\tilde{C}_i).$$

Lemma 20. *For any $N \in \mathbb{N}$,*

$$\underline{\tilde{q}}I \leq \mathbf{Q}_N \leq \bar{\tilde{q}}I,$$

with

$$\underline{\tilde{q}} = \frac{\underline{\tilde{\varepsilon}}^2}{\bar{\tilde{r}}}, \quad \bar{\tilde{r}} = \max_i \left\{ \left\| \tilde{R}_i \right\|, \left\| \tilde{R}_{i,i+1} \right\| \right\},$$
$$\bar{\tilde{q}} = \frac{\bar{\tilde{\varepsilon}}^2}{\underline{\tilde{r}}}, \quad \underline{\tilde{r}} = \min_i \left\{ \sigma_{\min}(\tilde{R}_i), \sigma_{\min}(\tilde{R}_{i,i+1}) \right\}.$$

Proof. Since $\mathbf{Q}_N = \mathbf{A}_N^T \mathbf{S}_N^{-1} \mathbf{A}_N$, from Lemma 19, it follows that

$$\frac{\underline{\tilde{\varepsilon}}^2}{\sigma_{\max}(\mathbf{S}_N)} I \le \mathbf{Q}_N \le \frac{\bar{\tilde{\varepsilon}}^2}{\sigma_{\min}(\mathbf{S}_N)} I.$$

The result then follows from $\underline{\tilde{r}} I \le \mathbf{S}_N \le \bar{\tilde{r}} I$.

Lemma 21. *For every* $1 \le k \le N$,

$$\underline{\tilde{q}} I \le \Delta_k, \Gamma_k(N), \Phi_k(N) \le \bar{\tilde{q}} I.$$

Our next goal is to bound the difference $\|\hat{x}_1(\mathcal{L}_{N+1}) - \hat{x}_1(\mathcal{L}_N)\|$. From (16) and (19), this requires an upper bound for $\left\| \Phi_j^{-1}(N+1) - \Phi_j^{-1}(N) \right\|$. This is given in the following lemma, whose proof also appears in Sui et al. (2018, Appendix).

Lemma 22. *For any* $1 \le j \le N$, *we have*

$$\left\| \Phi_j^{-1}(N+1) - \Phi_j^{-1}(N) \right\| \le \tilde{q}^{-1} \left(e^{\tilde{\psi}_N \tilde{\lambda}_N^{N-j}} - 1 \right)$$

with

$$\tilde{\psi}_N = \sqrt{\bar{n} |\mathcal{T}_N|} \bar{\tilde{\xi}}_N,$$
$$\bar{\tilde{\xi}}_N = \max_{i \le N} \log \sigma_{\max}[I + (\tilde{C}_{i,i+1}^T \tilde{R}_{i,i+1}^{-1} \tilde{C}_{i,i+1}) \cdot (\tilde{C}_i^T \tilde{R}_i^{-1} \tilde{C}_i)^{-1}],$$
$$\tilde{\lambda}_N = \frac{\tilde{\alpha}_{1,N}}{\tilde{\alpha}_{1,N} + \tilde{\beta}_{1,N}} \frac{\tilde{\alpha}_{2,N}}{\tilde{\alpha}_{2,N} + \tilde{\beta}_{2,N}},$$

where

$$\tilde{\alpha}_{1,N} = \max_{i \le N} \|\tilde{C}_{i,i+1}^T \tilde{R}_{i,i+1}^{-1} \tilde{C}_{i,i+1}\|,$$
$$\tilde{\alpha}_{2,N} = \max_{i \le N} \|\tilde{C}_{i+1,i} (\tilde{C}_{i+1}^T \tilde{R}_{i+1}^{-1} \tilde{C}_{i+1})^{-1} \tilde{C}_{i+1,i}^T\|,$$
$$\tilde{\beta}_{1,N} = \min_{i \le N} \sigma_{\min}(\tilde{C}_i^T \tilde{R}_i^{-1} \tilde{C}_i), \quad \tilde{\beta}_{2,N} = \min_{i \le N} \sigma_{\min}(\tilde{R}_{i,i+1}).$$

Combining the results in Lemmas 21–22 and (16), we can obtain upper bounds for $\left\| [\Sigma_{N+1}]_{1,j} - [\Sigma_N]_{1,j} \right\|$ and $\left\| [\Sigma_N]_{1,j} \right\|$. These are given in Lemmas 23 and 24, respectively.

Lemma 23. *For any* $1 \le j \le N$,

$$\left\| [\Sigma_{N+1}]_{1,j} - [\Sigma_N]_{1,j} \right\| \le \bar{\tilde{q}}^{-1} \tilde{r}^j \left(e^{\tilde{\psi}_N \tilde{\lambda}_N^{N-j}} - 1 \right)$$

with $\tilde{r} = \bar{\tilde{q}} / \underline{\tilde{q}}$.

Proof. From Sui et al. (2018, Lemma 31), for all $k \in \mathbb{N}$,

$$\left\| Q_{k,k+1} \right\| \le \bar{\tilde{q}}.$$

We then have

$$\begin{aligned}
&\left\| [\Sigma_{N+1}]_{1,j} - [\Sigma_N]_{1,j} \right\| \\
&= \left\| \left(\prod_{k=1}^{j-1} \Delta_k^{-1} Q_{k,k+1} \right) \left(\Phi_j^{-1}(N+1) - \Phi_j^{-1}(N) \right) \right\| \\
&\le \left(\prod_{k=1}^{j-1} \left\| \Delta_k^{-1} Q_{k,k+1} \right\| \right) \left\| \Phi_j^{-1}(N+1) - \Phi_j^{-1}(N) \right\| \\
&\le \bar{\tilde{q}}^{j-1} \left(\prod_{k=1}^{j-1} \left\| \Delta_k^{-1} \right\| \right) \left\| \Phi_j^{-1}(N+1) - \Phi_j^{-1}(N) \right\|.
\end{aligned}$$

Then, using Lemmas 21 and 22, we get

$$\begin{aligned}
\left\| [\Sigma_{N+1}]_{1,j} - [\Sigma_N]_{1,j} \right\| &\le \frac{1}{\underline{\tilde{q}}} \left(\frac{\bar{\tilde{q}}}{\underline{\tilde{q}}} \right)^{j-1} \left(e^{\tilde{\psi}_N \tilde{\lambda}_N^{N-j}} - 1 \right) \\
&\le \underline{\tilde{q}}^{-1} \tilde{r}^j \left(e^{\tilde{\psi}_N \tilde{\lambda}_N^{N-j}} - 1 \right).
\end{aligned}$$

Lemma 24. *For all* $1 \le j \le N$,

$$\left\| [\Sigma_N]_{1,j} \right\| \le \tilde{c} \tilde{\iota}^j,$$

with

$$\tilde{c} = \frac{\tilde{r} - 1}{2 \underline{\tilde{q}} \tilde{\iota}}, \quad \tilde{\iota} = \frac{\sqrt{\tilde{r}} - 1}{\sqrt{\tilde{r}} + 1}.$$

Proof. Since $\mathbf{Q}_N$ is 2-banded (Sui et al., 2018 Definition 32), the result it follows from Sui et al. (2018, Lemma 33), by letting $a = \underline{\tilde{q}}$ and $b = \bar{\tilde{q}}$.

It follows from (19) that, in addition to the bounds given in Lemmas 23 and 24, we also need an upper bound for $\|q_i\|$. This is given in the following lemma.

Lemma 25. *For any* $N \in \mathbb{N}$,

$$\max_{n \le N} \|q_n\| \le \tilde{\eta}_N,$$

with

$$\tilde{\eta}_N = \max_{i \le N} \frac{2^{3/2} \bar{\tilde{\varepsilon}} \sqrt{\bar{m} |\mathcal{T}_i|} \bar{\tilde{z}}_N}{\underline{\tilde{r}}},$$
$$\bar{\tilde{z}}_N = \max_{i \le N} \left\{ \|\tilde{z}_i\|_\infty, \|\tilde{z}_{i,i+1}\|_\infty \right\}.$$

Proof. From (14),

$$\|q_n\| \le \frac{\bar{\tilde{\varepsilon}} (\|y_n\| + \|y_{n-1}\|)}{\underline{\tilde{r}}}. \tag{20}$$

For any $n = 1, 2, \ldots, N$, the result then follows from

$$\|y_n\| \le \max_{i \le N} \sqrt{2 \bar{m} |\mathcal{T}_i|} \|y_n\|_\infty \le \max_{i \le N} \sqrt{2 \bar{m} |\mathcal{T}_i|} \bar{\tilde{z}}_N.$$

We now state the main result of this subsection.

Lemma 26. *For any* $1 \le J \le N$,

$$\begin{aligned}
&\left\| \hat{x}_1(\mathcal{L}_{N+1}) - \hat{x}_1(\mathcal{L}_N) \right\| \\
&\le \tilde{\eta}_{N+1} \left(\frac{\tilde{r}^J}{(\tilde{r} - 1) \underline{\tilde{q}}} \left(e^{\tilde{\psi}_N \tilde{\lambda}_N^{N-J}} - 1 \right) + \frac{2 \tilde{c}}{1 - \tilde{\iota}} \tilde{\iota}^J \right).
\end{aligned}$$

Proof. From Lemmas 23–25 and (19),

$$\begin{aligned}
&\left\| \hat{x}_1(\mathcal{L}_{N+1}) - \hat{x}_1(\mathcal{L}_N) \right\| \\
&\le \sum_{j=1}^{J-1} \left\| [\Sigma_{N+1}]_{1,j} - [\Sigma_N]_{1,j} \right\| \|q_j\| \\
&\quad + \sum_{j=J}^{N} \left\| [\Sigma_{N+1}]_{1,j} - [\Sigma_N]_{1,j} \right\| \|q_j\| \\
&\quad + \left\| [\Sigma_{N+1}]_{1,N+1} \right\| \|q_{N+1}\| \\
&\le \tilde{\eta}_{N+1} \left(\sum_{j=1}^{J-1} \bar{\tilde{q}}^{-1} \tilde{r}^j \left(e^{\tilde{\psi}_N \tilde{\lambda}_N^{N-j}} - 1 \right) + 2 \tilde{c} \sum_{j=J}^{N} \tilde{\iota}^j + \tilde{c} \tilde{\iota}^{N+1} \right) \\
&\le \tilde{\eta}_{N+1} \left(\sum_{j=1}^{J-1} \bar{\tilde{q}}^{-1} \tilde{r}^j \left(e^{\tilde{\psi}_N \tilde{\lambda}_N^{N-J}} - 1 \right) + 2 \tilde{c} \sum_{j=J}^{N+1} \tilde{\iota}^j \right)
\end{aligned}$$

$$=\tilde{\eta}_{N+1}\left(\frac{\tilde{\bar{q}}^{-1}\left(e^{\tilde{\psi}_N\tilde{\lambda}_N^{N-J}}-1\right)}{\tilde{r}-1}+2\tilde{c}\frac{\tilde{\iota}^J-\tilde{\iota}^{N+2}}{1-\tilde{\iota}}\right)$$
$$\leq\tilde{\eta}_{N+1}\left(\frac{\tilde{\bar{q}}^{-1}\left(e^{\tilde{\psi}_N\tilde{\lambda}_N^{N-J}}-1\right)}{\tilde{r}-1}+2\tilde{c}\frac{\tilde{\iota}^J}{1-\tilde{\iota}}\right).$$

6.2.2. Bound of the increment in an arbitrary graph

For each N, the value of $\hat{x}_1(N)$ obtained by running Algorithm 1 on $\mathcal{G}$ equals the one $\hat{x}_1(\mathcal{A}_N)$ obtained by running the same algorithm on the equivalent acyclic graph $\mathcal{A}_N$. The latter in turn equals the one $\hat{x}_1(\mathcal{L}_N)$ obtained by running the algorithm on the equivalent line graph $\mathcal{L}_N$. Then, from Lemma 26 we obtain the following result, which applies to an arbitrary graph $\mathcal{G}$, and whose proof appears in Sui et al. (2018, Appendix).

Lemma 27. *If $\check{\kappa} < 1$, then, in the notation of Theorem 17, for all $N \in \mathbb{N}$,*

$$\left\|\hat{x}_1(N+1)-\hat{x}_1(N)\right\| \leq \check{\bar{\chi}}\check{\kappa}^N,$$

where

$$\check{\bar{\chi}} = \frac{\check{\psi}\check{\eta}}{(\bar{q}-\underline{q})\lambda}+\frac{2\check{\eta}c}{1-\iota}, \quad \check{\kappa} = \max\{\bar{u}\sqrt{\lambda}, \sqrt{\bar{u}}\iota^{1/\check{\zeta}}\},$$

with

$$\check{\psi} = \left(e^{\bar{\xi}\sqrt{\bar{n}(\bar{u}+1)}}-1\right), \quad \check{\eta} = \frac{\bar{\varepsilon}\bar{z}\sqrt{8\bar{m}(\bar{u}+1)}}{\underline{r}},$$
$$\check{\zeta} = 2+\log_{\frac{1}{\sqrt{\lambda}}}(\bar{q}/\underline{q}), \quad c = \frac{\bar{q}-\underline{q}}{2\bar{q}\underline{q}\iota},$$
$$\bar{\xi} = \max_j \log\|I+(\sum_{k\in\mathcal{N}_j} C_{j,k}^T R_{j,k}^{-1}C_{j,k})\left(C_j^T R_j^{-1}C_j\right)^{-1}\|,$$
$$\bar{z} = \max_{i,j}\left\{\|z_i\|_\infty, \|z_{i,j}\|_\infty\right\}.$$

Remark 28. The above lemma shows that, when $\check{\kappa} < 1$, the sequence of state estimates produced by Algorithm 1 converges exponentially.

6.2.3. Accuracy analysis

Let $\hat{x}_1^{\mathrm{WLS}}(N)$ denote the centralized WLS estimate of x_1 obtained by considering only the subgraph $\mathcal{N}_1(N-1)$ (i.e., of nodes in $\mathcal{G}$ which are within $N-1$ steps away from node 1). It follows from Tai et al. (2013) that, in an acyclic graph, if Algorithm 1 is initialized by $\mathbb{Q}^0$ (as done in (3)), it generates the true WLS estimate. Hence, based on the definition of the loop-free depth l_1 (see Notation 10), for any $N \leq l_1+1$,

$$\hat{x}_1(N) = \hat{x}_1^{\mathrm{WLS}}(N). \tag{21}$$

From Lemma 27, if $\check{\kappa} < 1$, for any $N \in \mathbb{N}$,

$$\left\|\hat{x}_1(N+1)-\hat{x}_1(N)\right\| \leq \check{\bar{\chi}}\check{\kappa}^N. \tag{22}$$

Let $\check{\mathcal{T}}_k$ denote the set of nodes in $\mathcal{G}$ which are precisely $k-1$ steps away from node 1. We collect all the nodes in $\check{\mathcal{T}}_k$ and their inner connections into a single node. This yields a graph $\check{\mathcal{L}}$ having line topology, whose structure is given in Sui et al. (2018, Appendix H). The number of nodes in $\check{\mathcal{L}}$ is given by

$$r_1 = \max_j d_{1,j}+1,$$

where $d_{i,j}$ is the minimum distance from node i to node j in $\mathcal{G}$.

Recall that Lemma 27 provides a bound for the state estimate increments for Algorithm 1. If we follow the steps of that proof, but considering $\check{\mathcal{L}}$ in place of $\mathcal{L}_N$, we would arrive to the following result.

Lemma 29. *Recall the notations in Theorem 17 and Lemma 27, for all $N \in \mathbb{N}$, if $\kappa < 1$,*

$$\left\|\hat{x}_1^{\mathrm{WLS}}(N+1)-\hat{x}_1^{\mathrm{WLS}}(N)\right\| \leq \begin{cases}\bar{\chi}\kappa^N, & N \leq r_1,\\ 0, & N > r_1,\end{cases} \tag{23}$$

where

$$\bar{\chi} = \frac{\bar{\psi}\bar{\eta}}{(\bar{q}-\underline{q})\omega}+\frac{2\bar{\eta}c}{1-\iota}$$
$$\bar{\psi} = e^{\bar{\xi}(\bar{u}+1)\sqrt{\bar{n}}}-1, \quad \bar{\eta} = \bar{\varepsilon}\bar{z}(\bar{u}+1)\sqrt{8\bar{m}}\underline{r}^{-1}.$$

Combining (21)–(22) and Lemma 29, we can prove our main result.

Proof of Theorem 17. Clearly, $\check{\kappa} \leq \kappa < 1$. Hence, from (21)–(23)

$$\begin{aligned}&\|\hat{x}_1(N)-\hat{x}_1^{\mathrm{WLS}}\|\\ =&\|\hat{x}_1(N)-\hat{x}_1(l_1+1)+\hat{x}_1^{\mathrm{WLS}}(l_1+1)-\hat{x}_1^{\mathrm{WLS}}(r_1)\|\\ \leq&\|\hat{x}_1(l_1+1)-\hat{x}_1(N)\|+\|\hat{x}_1^{\mathrm{WLS}}(l_1+1)-\hat{x}_1^{\mathrm{WLS}}(r_1)\|\\ \leq&\check{\bar{\chi}}\sum_{t=l_1+1}^{N-1}\check{\kappa}^t+\bar{\chi}\sum_{t=l_1+1}^{r_1-1}\kappa^t \leq \frac{\bar{\chi}}{1-\kappa}\kappa^{l_1+1}+\frac{\check{\bar{\chi}}}{1-\check{\kappa}}\check{\kappa}^{l_1+1}.\end{aligned}$$

Since $a_1 \geq \alpha_1$, $a_2 \geq \alpha_2$, $b_1 \leq \beta_1$ and $b_2 \leq \beta_2$, it follows that $\omega \geq \lambda$ and $\zeta \geq \check{\zeta}$. We also have $\check{\bar{\chi}} \leq \bar{\chi}$ and $\check{\kappa} \leq \kappa$. It then follows that

$$\|\hat{x}_1(N)-\hat{x}_1^{\mathrm{WLS}}\| \leq \frac{2\bar{\chi}}{1-\kappa}\kappa^{l_1+1}.$$

Since the quantity $\frac{2\bar{\chi}}{1-\kappa}$ only depends on $\bar{m}, \bar{u}, \bar{n}$ and the system parameters $C_{i,j}, C_i, R_{i,j}, R_i, z_{i,j}, z_i$ for some i and j, the result then follows.

7. Simulations

In this section we present experimental evidence to support our claims, namely, that in the case of cyclic communication graphs, the studied distributed WLS algorithm (DWLS) converges faster than the iterative matrix inversion (IMI) algorithm in Marelli and Fu (2015), and that the accuracy of the DWLS algorithm at a given node improves with the size of the loop-free depth of that node. To this end, we use a network formed by 330 nodes, whose communication graph is depicted in Fig. 1. In this network, all nodes have the same measurements equations, which are given by

$$z_i = x_i + v_i;$$
$$z_{i,j} = 0.4x_i + 0.4x_j + v_{i,j}.$$

with $x_i \in \mathbb{R}^3$, $R_i = R_{i,j} = 0.01$, for all $i = 1, 2, \ldots, 330$ and $j \in \mathcal{N}_i$.

In the first simulation we compare the convergence rate of the DWLS and IMI methods. As explained in Marelli and Fu (2015), before starting with the matrix inversion iterations, the IMI method needs to invest a number δ_{IMI} of iterations in order to obtain estimates of the largest and smallest eigenvalues of certain matrix. This delayed start is required in order to avoid that the transients caused by too rough estimates of these eigenvalues bring the estimation mismatch (with respect to the estimation yielded by the centralized WLS method) to very big values from which the algorithm would take a long time to converge. In Fig. 2 we show the combined estimation mismatch of all 330 nodes, yield by the DWLS algorithm and the IMI algorithm with δ_{IMI} ranging from 0 to 6. We see that the DWLS algorithm converges much faster than the IMI one, regardless of the value of δ_{IMI} used in the latter.

In the second simulation we evaluate the mismatches between the estimation and its associated covariance, produced at each node, and at time l_i+1 (recall that l_i denotes the loop-free depth

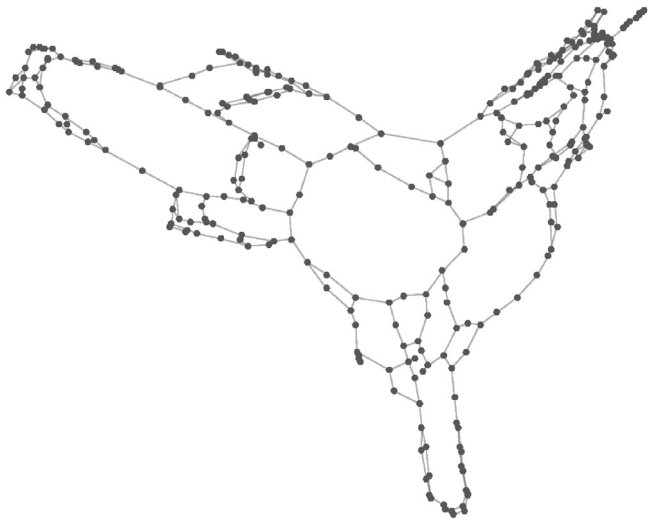

Fig. 1. Network communication graph.

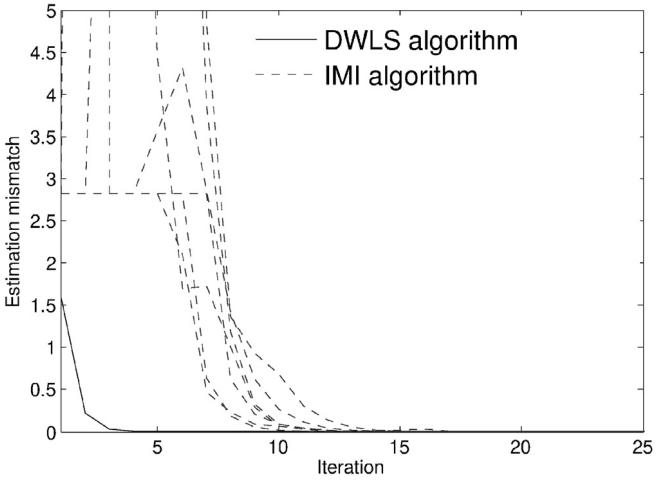

Fig. 2. Comparison between DWLS and IMI algorithms. The delayed start δ_{IMI} of the IMI algorithm ranges from 0 to 6.

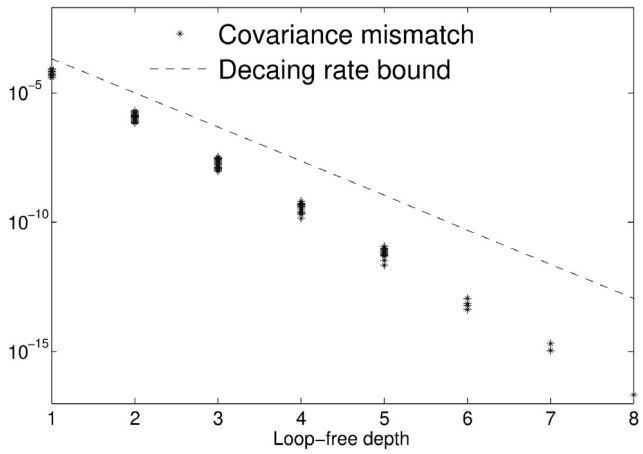

Fig. 3. Covariance mismatch between DWLS and centralized WLS.

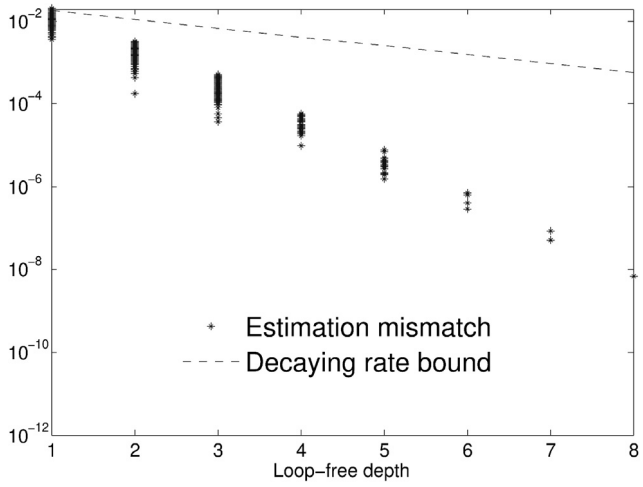

Fig. 4. Estimation mismatch between DWLS and centralized WLS.

of node i), with respect to those yielded by the centralized WLS method. Figs. 3 and 4 show these differences for the covariance and estimate, respectively, for each node, as a function of the loop-free depth. We see how both differences decay exponentially with the loop-free depths of each node. We also show in the same figures the bound on these decays derived in Theorems 11 and 17, respectively.

8. Conclusions

A recently proposed distributed WLS estimation algorithm converges in finite time if the communication graph is acyclic. We studied the accuracy of this algorithm, when used in cyclic graphs. We showed that, for a class of systems satisfying certain requirements in terms of topological sparsity and signal-to-noise ratio, the error between the state estimate yielded by this distributed algorithm, and that from centralized WLS, decreases exponentially at each node, with the increase of its local loop-free depth. The same property holds for the difference between the estimation error covariance produced by the distributed algorithm and that from centralized WLS. The derived expressions are explicit and easy to interpret. An implication of our results is that, even in applications where the communication graph has a cyclic topology, due to its faster convergence, the studied algorithm may be a preferred option over algorithms based on iterative matrix inversion, provided that the loop-free depths of those nodes of interest are sufficiently large.

References

Acikmese, B., Mandić, M., & Speyer, J. (2014). Decentralized observers with consensus filters for distributed discrete-time linear systems. *Automatica*, *50*(4), 1037–1052.

Ajgl, J., & Šimandl, M. (2014). On linear estimation fusion under unknown correlations of estimator errors. In *IFAC WC*.

Battistelli, G., & Chisci, L. (2014). Kullback-Leibler average, consensus on probability densities, and distributed state estimation with guaranteed stability. *Automatica*, *50*(3), 707–718.

Bougerol, P. (1993). Kalman filtering with random coefficients and contractions. *SIAM Journal on Control and Optimization*, *31*(4), 942–959.

Calafiore, G., & Abrate, F. (2009). Distributed linear estimation over sensor networks. *International Journal of Control*, *82*(5), 868–882.

Cattivelli, F., & Sayed, A. (2010). Diffusion strategies for distributed Kalman filtering and smoothing. *IEEE Transactions on Automatic Control*, *55*(9), 2069–2084.

Chen, L., Arambel, P., & Mehra, R. (2002). Estimation under unknown correlation: covariance intersection revisited. *IEEE Transactions on Automatic Control*, *47*(11), 1879–1882.

Farina, M., Ferrari-Trecate, G., & Scattolini, R. (2010). Moving-horizon partition-based state estimation of large-scale systems. *Automatica*, *46*(5), 910–918.

Garin, F., & Schenato, L. (2010). A survey on distributed estimation and control applications using linear consensus algorithms. *Lecture Notes Contributions Infectious*, *406*(1), 75–107.
Gupta, V., Dana, A., Hespanha, J., Murray, R., & Hassibi, B. (2009). Data transmission over networks for estimation and control. *IEEE Transactions on Automatic Control*, *54*(8), 1807–1819.
Haber, A., & Verhaegen, M. (2013). Moving horizon estimation for large-scale interconnected systems. *IEEE Transactions on Automatic Control*, *58*(11), 2834–2847.
Ihler, A., Iii, J., & Willsky, A. (2005). Loopy belief propagation: convergence and effects of message errors. *Journal of Machinery Learning Research*, 905–936.
Khan, U., & Moura, J. (2008). Distributing the Kalman filter for large-scale systems. *IEEE Transactions on Signal Processing*, *56*(10), 4919–4935.
Li, J., & Alregib, C. (2009). Distributed estimation in energy-constrained wireless sensor networks. *IEEE Transactions on Signal Processing*, *57*(10), 3746–3758.
Marelli, D., & Fu, M. (2015). Distributed weighted least-squares estimation with fast convergence for large-scale systems. *Automatica*, *51*, 27–39.
Matei, I., & Baras, J. (2012). Consensus-based linear distributed filtering. *Automatica*, *48*(8), 1076–1082.
Meurant, G. (1992). A review on the inverse of symmetric tridiagonal and block tridiagonal matrices. *SIAM Journal on Matrix Analysis and Applications*, *13*(3), 707–728.
Olfati-Saber, R. (2005). Distributed Kalman filter with embedded consensus filters. *CDC*, *1*, 8179–8184.
Pasqualetti, F., Carli, R., & Bullo, F. (2012). Distributed estimation via iterative projections with application to power network monitoring. *Automatica*, *48*(5), 747–758.
Ribeiro, A., & Giannakis, G. (2006a). Bandwidth-constrained distributed estimation for wireless sensor networks-part I: Gaussian case. *IEEE Transactions on Signal Processing*, *54*(3), 1132–1143.
Ribeiro, A., & Giannakis, G. (2006b). Bandwidth-constrained distributed estimation for wireless sensor networks-part II: unknown probability density function. *IEEE Transactions on Signal Processing*, *54*(7), 2784–2796.
Sui, Tianju, Marelli, Damin, Fu, Minyue, & Lu, Renquan (2018). Accuracy analysis for distributed weighted least-squares estimation in finite steps and loopy networks, arXiv:1806.09104[cs.MA].
Tai, X., Lin, Z., Fu, M., & Sun, Y. (2013). A new distributed state estimation technique for power Networks. In *ACC* (pp. 3338–3343).
Tatikonda, S. (2003). Convergence of the sum-product algorithm. In *ITW* (pp. 222–225).
Tatikonda, S., & Jordan, M. (2002). Loopy belief propagation and Gibbs measures. In *CUAI* (pp. 493–500).
Weiss, Y. (2000). Correctness of local probability propagation in graphical models with loops. *Neural Computation*, *12*, 1–41.
Zhou, T. (2013). Coordinated one-step optimal distributed state prediction for a networked dynamical system. *IEEE Transactions on Automatic Control*, *58*(11), 2756–2771.
Zhou, T. (2015). On the controllability and observability of networked dynamic systems. *Automatica*, *52*, 63–75.

Tianju Sui received B.Eng. and Ph.D. degrees from Zhejiang University, Hangzhou, China, in 2012 and 2017 respectively. He is currently working in Dalian University of Technology. His main research area includes networked estimation, distributed estimation and CPS security.

Damián Marelli received his Bachelor's Degree in Electronics Engineering from the Universidad Nacional de Rosario, Argentina in 1995. He also received his Bachelor (Honous) degree in Mathematics and Ph.D. degree in Electrical Engineering, both from the University of Newcastle, Australia in 2003. From 2004 to 2005 he was postdoc at the Laboratoire d'Analyse Topologie et Probabilits, CNRS/Université de Provence, France. From 2005 to 2015 he was Research Academic at the Centre for Complex Dynamic Systems and Control, the University of Newcastle, Australia. In 2007 he received a Marie Curie Postdoctoral Fellowship, hosted at the University of Vienna, and in 2011 he received a Lise Meitner Senior Fellowship, hosted at the Austrian Academy of Sciences. Since 2016, he is Professor at the School of Automation, Guangdong University of Technology, China and Independent Researcher appointment at the French-Argentinean International Center for Information and Systems Sciences, National Scientific and Technical Research Council, Argentina. His main research interests include system theory, statistical signal processing and distributed processing.

Minyue Fu received his Bachelor's Degree in Electrical Engineering from the University of Science and Technology of China, Hefei, China, in 1982, and M.S. and Ph.D. degrees in Electrical Engineering from the University of Wisconsin-Madison in 1983 and 1987, respectively. From 1987 to 1989, he served as an Assistant Professor in the Department of Electrical and Computer Engineering, Wayne State University, Detroit, Michigan. For the summer of 1989, he was employed by the Universite Catholoque de Louvain, Belgium, as a Maitre de Conferences Invited. He joined the Department of Electrical and Computer Engineering, the University of Newcastle, Australia, in 1989 and was promoted to a Chair Professor in Electrical Engineering in 2002. He has served as the Head of Department for Electrical and Computer Engineering and Head of School of Electrical Engineering and Computer Science over a period of 7 years. In addition, he was a Visiting Associate Professor at University of Iowa in 1995–1996, a Visiting Professor at Nanyang Technological University, Singapore, 2002, and Visiting Professor at Tokyo University in 2003. He has held a ChangJiang Visiting Professorship at Shandong University, a visiting Professorship at South China University of Technology, and a Qian-ren Professorship at Zhejiang University in China. He was elected to Fellow of IEEE in late 2003. His main research interests include control systems, signal processing and communications. His current research projects include networked control systems, multi-agent systems, smart electricity networks and super-precision positioning control systems. He has been an Associate Editor for the IEEE Transactions on Automatic Control, IEEE Transactions on Signal Processing, Automatica, and Journal of Optimization and Engineering.

Renquan Lu received his Ph.D. degree in Control Science and Engineering from Zhejiang University, Hangzhou, China, in 2004. He was supported by the National Science Fund for Distinguished Young Scientists of China in 2014, honored as the Distinguished Professor of Pearl River Scholars Program of Guangdong Province, the Distinguished Professor of Yangtze River Scholars Program by the Ministry of Education of China in 2015 and 2017, respectively. Currently, he is a professor of the School of Automation at Guangdong University of Technology Guangzhou China. His research interests include complex systems, networked control systems, and nonlinear systems.