2018 IEEE 14th International Conference on Control and Automation (ICCA)
June 12-15, 2018. Anchorage, Alaska, USA

Distributed Estimation in Networks of Linear Time-invariant Systems

Mohsen Zamani, Damian Marelli, Brett Ninness and Minyue Fu.

Abstract— **This paper is concerned with the problem of distributed Kalman filtering in a network of several interconnected subsystems. We consider networks, which can be either homogeneous or heterogeneous, of linear time-invariant subsystems, given in state-space form. We propose a distributed Kalman filtering scheme for this setup. The proposed scheme provides estimates based only on locally available measurements. We compare its outcomes with those of a centralized Kalman filter, which offers the best minimum error variance estimate, using all measurements available all over the network. We show that the estimate produced by the proposed method asymptotically approaches to that of the centralized Kalman filter, i.e., the optimal one with global knowledge of all network parameters, and we are able to bound the convergence rate. Moreover, if the initial states of all subsystems are mutually uncorrelated, the estimates of these two schemes are identical at each time step.**

I. Introduction

There has been an increasing effort in the study of distributed estimation in a network environment. This is due to its broad applications in many areas, including formation control [1], [2], distributed sensor network [3] and cyber security [4], [5]. This paper examines the problem of distributed estimation in a network of subsystems represented by a finite dimensional state-space model. In particualr, our focus is on networks of finite-dimensional linear discrete-time dynamical systems that arise through static interconnections of a finite number of such systems. Such models arise naturally in applications of linear networked systems, e.g., for cyclic pursuit [6]; shortening flows in image processing [7], or for the discretization of partial differential equations [8]. This framework can be seen as a generalization of consensus algorithm, widely studied in the literature [9], [10]. Hence, results of the current paper are applicable to consensus-based networks as well as to more general types of networks modelled with static interconnection links. We suppose that states of each subsystem are manipulated by two main components, namely, a local term and a term associated with the neighboring subsystems. Furthermore, there exist noise components that may affect the states and measurements. It is worth noting that such a scenario arises in different applications like security, where normally the estimation of states is required to calculate a residue for attack detection [4]; and, in formation control [1], [11]–[13], where each subsystem integrates measurements from its nearby subsystems, and states of each subsystem need to be estimated for distributed control design purposes. The main objective of this paper is to collectively estimate the states of all subsystems within such a network. To this end, we propose a novel distributed version of the celebrated Kalman filter.

The current paper, in broad sense, belongs to the large body of literature regarding distributed estimation. One can refer to [14]–[24] and the survey paper [25], as well as references listed therein, for different variations of distributed estimation methods among a group of subsystems within a network. A consensus based Kalman filter was proposed in [19]. The author of [20] utilized a linear matrix inequality to minimize a H_∞ index associated with a consensus based estimator, which can be implemented locally. Some of the results there were then extended to the case of switching topology in [21]. The same problem was solved using the minimum energy filtering approach in [22]. A common drawback of the state estimation methods described above is that, being based on consensus, they require, in theory, an infinite number of consensus iterations at each time step. This results in computational and communication overload. To avoid this, in this paper we exploit the network structure to achieve a distributed Kalman filter method which requires only one prediction/update step at each time step. We remark that in the current paper each subsystem is dedicated to calculate its own states; however, in the above-mentioned works, all subsystems contribute toward estimation of a common global state.

To show the effectiveness of the proposed algorithm, we compare our method with the classical (centralized) Kalman filter, which is known to be optimal (in the minimum error covariance sense). The classical method requires the simultaneous knowledge of parameters and measurements from all subsystems within the network to carry out the estimation. In contrast, our proposed distributed estimation algorithm runs a local Kalman filter at each subsystem, which only requires the knowledge of local measurements and parameters, as well as measurements from neighbor subsystems. Hence, it can be implemented in a fully distributed fashion. We show that the state estimate, and its associated estimation error covariance matrix, produced by the proposed distributed method asymptotically converge to those produced by the centralized Kalman filter. We provide bounds for the convergence of both the estimate and the estimation error covariance matrix. We further establish that when initial the states of all subsystems are mutually uncorrelated, the estimates produced by both

M. Zamani, B. Ninness and M. Fu are with the School of Electrical Engineering and Computer Science, University of Newcastle, NSW 2308, Australia. M. Fu and D. Marelli are with Science and Engineering and State Key Laboratory of Industrial Control Technology, Zhejiang University, 388 Yuhangtang Road Hangzhou, Zhejiang Province, 310058, P. R. China. D. Marelli is also with the CIFASIS/CONICET, 2000 Rosario, Argentina. e-mail: {mohsen.zamani, damian.marelli, brett.ninness, minyue.fu}@newcastle.edu.au.

978-1-5386-6089-8/18/$31.00 ©2018 IEEE

methods i.e. centralized and distributed, are identical.

The rest of the paper is structured as follows. In Section II, we describe the network setup and its associated centralized Kalman filter. In Section IV, we describe the proposed distributed Kalman filter scheme. In Section V, we demonstrate the asymptotic equivalence between the proposed distributed filter and the centralized one, and provide bounds for the convergence of the estimates and their associated estimation error covariances. Simulation results that support our theoretical claims are presented in Section VI. Finally, concluding remarks are given in Section VII.

II. System Description

In this paper we study networks of N time-invariant subsystems. Subsystem i is represented by the following state-space model

$$x_{k+1}^{(i)} = A^{(i)}x_k^{(i)} + z_k^{(i)} + w_k^{(i)}, \quad (1)$$

$$y_k^{(i)} = C^{(i)}x_k^{(i)} + v_k^{(i)}. \quad (2)$$

The subsystems are interconnected as follows

$$z_k^{(i)} = \sum_{j\in\mathcal{N}_i} L^{(i,j)}y_k^{(j)}, \quad (3)$$

where $x_k^{(i)} \in \mathbb{R}^{n_i}$ is the state, $y_k^{(i)} \in \mathbb{R}^{p_i}$ the output, $w_k^{(i)}$ is an i.i.d Gaussian disturbance process with $w_k^{(i)} \sim \mathcal{N}(0, Q_i)$, and $v_k^{(i)}$ is an i.i.d. Gaussian measurement noise process with $v_k^{(i)} \sim \mathcal{N}(0, R_i)$. We further suppose that $\mathcal{E}\left(w_k^{(i)}w_k^{(j)^\top}\right) = 0$ and $\mathcal{E}\left(v_k^{(i)}v_k^{(j)^\top}\right) = 0$, $\forall i \neq j$ and $\mathcal{E}\left(x_k^{(i)}w_k^{(j)^\top}\right) = 0$, $\mathcal{E}\left(x_k^{(i)}v_k^{(j)^\top}\right) = 0$ $\forall i, j$. We also denote the neighbor set of the subsystem i by $\mathcal{N}_i = \left\{j : L^{(i,j)} \neq 0\right\}$.

Remark 1: We note in (1)-(2) that the coupling between neighboring subsystems is solely caused through the $z_k^{(i)}$ term in‘(3). The main motivation for considering such coupling comes from distributed control, where‘(1) represents the model of an autonomous subsystem (or agent) with $z_k^{(i)}$ being the control input, and (3) represents a distributed control protocol, which employs feedback only from neighboring measurements. This type of distributed control is not only common for control of multi-agent systems (see, for example, [2], [11]–[13]), but also realistic for large networked systems, since only neighbouring information is both easily accessible and most useful for each subsystem.

Notice that the dynamical descriptions (1)-(3) can be regarded as a very general setting for the well-known consensus algorithm [9], i.e., when it is run over a group of interconnected multi-input-multi-output linear subsystems expressed in state space form. Additionally, this model can capture interactions within linear dynamical networks. Interested readers can refer to [26], [27], [28] and [29], where the authors exploited a similar model for conducting system identification analysis in linear dynamical networks. Finally, this model turns out to be an effective one for studying peroperties of networked subsystems [5].

We emphasize that the distributed state estimation problem arises for the networked system (1)-(3) because of our allowance for measurement noises $v_k^{(i)}$ in (2). This consideration is very important for applications because measurement noises are unavoidable in practice. This also sharply distinguishes our distributed control formulation from most distributed control algorithms in the literature, where perfect state measurement is often implicitly assumed.

We define $\xi_k^\top = \left[\left(\xi_k^{(1)}\right)^\top, \cdots, \left(\xi_k^{(I)}\right)^\top\right]$ and $\Xi_k = \{\xi_1, \cdots, \xi_k\}$, where (ξ, Ξ) stands for either (x, X), (y, Y), (z, Z), (w, W) or (v, V); moreover, we denote $\Upsilon = \operatorname{diag}\left\{\Upsilon^{(1)}, \cdots, \Upsilon^{(I)}\right\}$, where Υ stands for either A, B, C, Q or R, and $L = \left[L^{(i,j)} : i, j = 1, \cdots, N\right]$.

Using the above notation, we let the initial state of all subsystems have the joint distribution $x_0 \sim \mathcal{N}(\mu, P)$. We can also write the aggregated model of the whole network as

$$\begin{aligned} x_{k+1} &= Ax_k + LCx_k + w_k + BLv_k \\ &= \tilde{A}x_k + e_k, \quad (4) \\ y_k &= Cx_k + v_k, \quad (5) \end{aligned}$$

with

$$\tilde{A} = A + LC \quad \text{and} \quad e_k = w_k + Lv_k. \quad (6)$$

It then follows that

$$\operatorname{cov}\left(\begin{bmatrix} e_k \\ v_k \end{bmatrix}\begin{bmatrix} e_k^\top & v_k^\top \end{bmatrix}\right) = \begin{bmatrix} \tilde{Q} & \tilde{S} \\ \tilde{S}^\top & R \end{bmatrix}, \quad (7)$$

where $\tilde{Q} = Q + LRL^\top$ and $\tilde{S} = LR$.

III. Centralized Kalman Filter

Consider the standard (centralized) Kalman filter. For all $k, l \in \mathbb{N}$, let

$$\begin{aligned} \hat{x}_{k|l} &\triangleq \mathcal{E}\left(x_k|Y_l\right), \\ \Sigma_{k|l} &\triangleq \mathcal{E}\left(\left[x_k - \hat{x}_{k|l}\right]\left[x_k - \hat{x}_{k|l}\right]^\top\right). \end{aligned} \quad (8)$$

Our aim in this subsection is to compute $\hat{x}_{k|k}$ in a standard centralized way. Notice that equation (7) implies that, in the aggregated system formed by (1)-(2), the process noise e_k and the measurement noise v_k are mutually correlated. Taking this into account, it follows from [30, S 5.5] that the prediction and update steps of the (centralized) Kalman filter are initialized by $\hat{x}_{0|0} = \mu$ and $\Sigma_{0|0} = P$, and proceed as follows:

1) **Prediction:**

$$\begin{aligned} \hat{x}_{k+1|k} &= \left(\tilde{A} - \tilde{S}R^{-1}C\right)\hat{x}_{k|k} + \tilde{S}R^{-1}y_k \\ &= A\hat{x}_{k|k} + Ly_k, \end{aligned} \quad (9)$$

and

$$\begin{aligned} \Sigma_{k+1|k} &= \left(M - \tilde{S}R^{-1}C\right)\Sigma_{k|k}\left(M - \tilde{S}R^{-1}C\right)^\top \\ &\quad + \tilde{Q} - \tilde{S}R^{-1}\tilde{S} \\ &= A\Sigma_{k|k}A^\top + Q. \end{aligned} \quad (10)$$

2) **Update:**

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k \left(y_k - C\hat{x}_{k|k-1}\right), \quad (11)$$
$$\Sigma_{k|k} = \left(I - K_k C\right)\Sigma_{k|k-1}, \quad (12)$$

with

$$K_k = \Sigma_{k|k-1}C^\top \left(C\Sigma_{k|k-1}C^\top + R\right)^{-1}. \quad (13)$$

IV. Distributed Kalman Filter

Consider the i-th subsystem (1)-(2). Notice that, since the measurements $y_k^{(j)}$, $j \in \mathcal{N}_i$, are known by the i-th subsystem, they can be treated as external inputs. This observation leads us to the following intuitive approach for a distributed Kalman filter scheme.

Let, for all $i = 1, \cdots, I$ and $k, l \in \mathbb{N}$,

$$\begin{aligned}\hat{x}_{k|l}^{(i)} &\triangleq \mathcal{E}\left(x_k^{(i)} | y_m^{(j)}; j \in \mathcal{N}_i \cup \{i\}, m = 1, \cdots, l\right),\\ \Sigma_{k|l}^{(i)} &\triangleq \mathcal{E}\left(\left[x_k^{(i)} - \hat{x}_{k|l}^{(i)}\right]\left[x_k^{(i)} - \hat{x}_{k|l}^{(i)}\right]^\top\right).\end{aligned} \quad (14)$$

Then, the prediction and update steps for the proposed distributed Kalman filter are initialized by $\hat{x}_{0|0}^{(i)} = \mu^{(i)}$ and $\Sigma_{0|0}^{(i)} = P^{(i,i)}$, and proceed as follows:

1) **Prediction:**

$$\hat{x}_{k+1|k}^{(i)} = A^{(i)}\hat{x}_{k|k}^{(i)} + \sum_{j\in\mathcal{N}_i} L_{i,j}^{(i,j)} y_k^{(j)}, \quad (15)$$
$$\Sigma_{k+1|k}^{(i)} = A^{(i)}\Sigma_{k|k}^{(i)}{A^{(i)}}^\top + Q^{(i)}. \quad (16)$$

2) **Update:**

$$\hat{x}_{k|k}^{(i)} = \hat{x}_{k|k-1}^{(i)} + K_k^{(i)}\left(y_k^{(i)} - C^{(i)}\hat{x}_{k|k-1}^{(i)}\right), \quad (17)$$
$$\Sigma_{k|k}^{(i)} = \left(I - K_k^{(i)}C^{(i)}\right)\Sigma_{k|k-1}^{(i)}, \quad (18)$$

with

$$K_k^{(i)} = \Sigma_{k|k-1}^{(i)}{C^{(i)}}^\top \left(C^{(i)}\Sigma_{k|k-1}^{(i)}{C^{(i)}}^\top + R^{(i)}\right)^{-1}. \quad (19)$$

V. Optimality analysis

Since the distributed Kalman filter approach given in Section IV is motivated by intuition, the question naturally arises as to which extent it is optimal. In this section we address this question. To this end, we define $\left(\hat{x}_{k|l}^\star, \Sigma_{k|l}^\star\right)$, where $\hat{x}_{k|l}^{\star\top} = \left[\hat{x}_{k|l}^{(i)\top} : i = 1, \cdots, N\right]$ and $\Sigma_{k|l}^\star = \operatorname{diag}\left(\Sigma_{k|l}^{(i)} : i = 1, \cdots, N\right)$, to be the outcomes of distributed filter and $\left(\hat{x}_{k|l}, \Sigma_{k|l}\right)$ to be those of centralized one. In Section V-A, we show that the estimation error covariance of the distributed filter $\Sigma_{k|k}^\star$ converges to that of the centralized one $\Sigma_{k|k}$, and provide a bound for this convergence. In Section V-B, we do the same for the convergence of $\hat{x}_{k|k}^\star$ to $\hat{x}_{k|k}$.

A. Convergence of $\Sigma_{k|k}^\star$ to $\Sigma_{k|k}$

In this section, we show that the covariance matrices $\Sigma_{k|k}$ and $\Sigma_{k|k}^\star$ exponentially converge to each other, and introduce a bound on $\left\|\Sigma_{k|k} - \Sigma_{k|k}^\star\right\|$. To this end, we require the following definition from [31, Def 1.4].

Definition 1: For $n \times n$ matrices $P, Q > 0$, the Riemannian distance is defined by

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

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

Several properties of the above definition, which we use to derive our results, are given in the following proposition.

Proposition 1: [32, Proposition 5] For $n \times n$ matrices $P, Q > 0$, the following holds true:

1) $\delta(P, P) = 0$.
2) $\delta\left(P^{-1}, Q^{-1}\right) = \delta\left(Q, P\right) = \delta\left(P, Q\right)$.
3) For any $m \times m$ matrix $W > 0$ and $m \times n$ matrix B, we have
$$\delta(W + BPB^\top, W + BQB^\top) \leq \frac{\alpha}{\alpha + \beta}\delta(P, Q),$$
where $\alpha = \max\{\|BPB^\top\|, \|BQB^\top\|\}$ and $\beta = \sigma_{\min}\left(W\right)$.
4) If $P > Q$, then $\|P - Q\| \leq \left(e^{\delta(P,Q)} - 1\right)\|Q\|$.

The main result of this section is given in Theorem 1. Its proof requires the following technical result. Its proof is omitted in this conference version, and is provided in [33].

Lemma 1: Let $\Gamma_{k|l} = \Sigma_{k|l}^{-1}$ and $\Gamma_{k|l}^\star = {\Sigma_{k|l}^\star}^{-1}$. Then

$$\begin{aligned}\left\|\Sigma_{k|k}\right\|, \left\|\Sigma_{k|k}^\star\right\| &\leq \sigma,\\ \left\|\Gamma_{k|k}\right\|, \left\|\Gamma_{k|k}^\star\right\| &\leq \omega,\end{aligned}$$

and

$$\delta\left(\Sigma_{k|k}, \Sigma_{k|k}^\star\right) \leq \upsilon^k \delta\left(P, P^\star\right), \quad (20)$$
$$\delta\left(\Gamma_{k|k}, \Gamma_{k|k}^\star\right) \leq \upsilon^k \delta\left(P, P^\star\right), \quad (21)$$

where

$$\sigma = \max\left\{\|P\|, \|P^\star\|, \left\|\bar{\Sigma}\right\|\right\}, \quad (22)$$
$$\omega = \max\left\{\left\|P^{-1}\right\|, \left\|{P^\star}^{-1}\right\|, \left\|\bar{\Sigma}^{-1}\right\|\right\}, \quad (23)$$

with $P^\star$ denoting the diagonal matrix formed by the block diagonal entries of the matrix P,

$$\upsilon = \upsilon_1\upsilon_2, \qquad \upsilon_1 = \frac{\sigma\|A\|^2}{\sigma\|A\|^2 + \|Q^{-1}\|^{-1}}, \quad (24)$$
$$\upsilon_2 = \frac{\omega}{\omega + \|U^{-1}\|^{-1}}, \qquad U = C^\top R^{-1} C,$$

and $\bar{\Sigma} = \lim_{k\to\infty}\Sigma_{k|k}$.

We now introduce the main result of the section, stating a bound on $\left\|\Sigma_{k|k} - \Sigma_{k|k}^\star\right\|$. Its proof is also omitted and provided in [33].

Theorem 1: Let $\tilde{\Sigma}_{k|l} = \Sigma_{k|l} - \Sigma^{\star}_{k|l}$ and $\tilde{\Gamma}_{k|l} = \Gamma_{k|l} - \Gamma^{\star}_{k|l}$. Then

$$\left\|\tilde{\Sigma}_{k|k}\right\| \leq \kappa\sigma\upsilon^k \quad \text{and} \quad \left\|\tilde{\Gamma}_{k|k}\right\| \leq \kappa\omega\upsilon^k,$$

where

$$\kappa = e^{\delta(P,P^{\star})} - 1.$$

B. Convergence of $\hat{x}^{\star}_{k|k}$ to $\hat{x}_{k|k}$

In this subsection, we study the convergence of state estimate $\hat{x}^{\star}_{k|k}$, obtained through the distributed method, and that of the centralized one $\hat{x}_{k|k}$. Moreover, we derive a bound on the error $\hat{x}_{k|k}$-$\hat{x}^{\star}_{k|k}$. We start by introducing two lemmas which are instrumental for establishing our main results. Their proofs, as well as that of our main result, are omitted in this conference version and provided in [33].

Lemma 2: Let $\tilde{x}_{k|l} = \hat{x}_{k|l} - \hat{x}^{\star}_{k|l}$. Then

$$\tilde{x}_{k+1|k} = H_k\tilde{x}_{k|k-1} + \xi_k. \tag{25}$$

where

$$\begin{aligned} H_k &= A\left(I - \Sigma_{k|k}U\right), \\ \xi_k &= a_k + b_k, \\ a_k &= A\Sigma_{k|k}\tilde{\Gamma}_{k|k}\hat{x}^{\star}_{k|k-1}, \\ b_k &= A\tilde{\Sigma}_{k|k}\Gamma^{\star}_{k|k}\hat{x}^{\star}_{k|k}. \end{aligned}$$

Lemma 3: Let

$$\Delta_k = \mathcal{E}\left(\tilde{x}_{k|k-1}\tilde{x}^{\top}_{k|k-1}\right). \tag{26}$$

Then

$$\Delta_k \leq H_k\Delta_{k-1}H_k^{\top} + \lambda\upsilon^k I, \tag{27}$$

where I is the identity matrix, υ is defined in (24), and

$$\lambda \triangleq \sup_{k\in\mathbb{N}}\left(\zeta + 2\sqrt{\zeta\left\|H_k\right\|^2\left\|\Delta_{k-1}\right\|}\right) < \infty, \tag{28}$$

with

$$\begin{aligned} \zeta &= (\alpha+\beta) + 2\sqrt{\alpha\beta}, \\ \alpha &= \kappa^2\omega^2\sigma^2\left\|A\right\|^2\left(\sigma\left\|A\right\|^2 + \left\|Q\right\|\right), \\ \beta &= \kappa^2\omega^2\sigma^3\left\|A\right\|^2. \end{aligned} \tag{29}$$

The following result states a family of upper bounds on the norm of the covariance matrix of $\tilde{x}_{k|l}$.

Theorem 2: Consider Δ_k as defined in (26). Let $H_k = V_kJ_kV_k{}^{-1}$ and $\bar{H} = \bar{V}\bar{J}\bar{V}^{-1}$ be the Jordan decompositions of H_k and $\bar{H}$, respectively. Then for every $\epsilon > 1$, there exists $k_\epsilon \in \mathbb{N}$ such that

$$\left\|\Delta_k\right\| \leq A_\epsilon\psi_\epsilon^k + B_\epsilon\upsilon^k,$$

where

$$A_\epsilon = \frac{\lambda\psi_\epsilon\phi_\epsilon}{\psi_\epsilon - \upsilon}, \qquad B_\epsilon = \frac{\lambda\upsilon\phi_\epsilon}{\upsilon - \psi_\epsilon}.$$

and

$$\begin{aligned} \psi_\epsilon &= \epsilon\rho\left(\bar{H}\right), \qquad \bar{H} = \lim_{k\to\infty}H_k, \\ \phi_\epsilon &= \epsilon\left\|\bar{V}\right\|^2\left\|\bar{V}^{-1}\right\|^2\left(\frac{m_\epsilon}{\epsilon\rho\left(\bar{H}\right)}\right)^{2(k_\epsilon-1)}, \\ m_\epsilon &= \max\left\{1, \left\|H_1\right\|, \cdots, \left\|H_{k_\epsilon-1}\right\|\right\}. \end{aligned} \tag{30}$$

Theorem 2 states that the covariance of the difference between $\hat{x}^{\star}_{k|k-1}$ and $\hat{x}_{k|k-1}$ is bounded by two exponential terms. The term $B_\epsilon\upsilon^k$ is due to the convergence of the Kalman gain $K^{\star}_k$ to K_k, while the term $A_\epsilon\psi^k_\epsilon$ is due to the convergence of the states given by the system dynamics. In order to use this result to show the asymptotic convergence of $\hat{x}^{\star}_{k|k-1}$ to $\hat{x}_{k|k-1}$, we need that $\upsilon < 1$ and $\psi_\epsilon < 1$, for some $\epsilon > 0$. While it is clear from (24) that the former is true, guaranteeing the latter is not that straightforward. The following proposition addresses this issue. Its proof appears in [33].

Proposition 2: If the pair $[A, C]$ is completely detectable and the pair $\left[A, Q^{1/2}\right]$ is completely stabilizable, then $\rho\left(\bar{H}\right) < 1$, where $\rho(\bar{H})$ denotes the spectral radius of matrix $\bar{H}$.

C. The case when the initial covariance is block diagonal

It turns out that, when the initial covariance matrix has a block diagonal structure both estimation methods are completely identical. This is summarized in the following corollary.

Corollary 1: Consider the network of subsystems (1)-(2). If the matrix P is block diagonal, then the distributed Kalman filter scheme (15)-(19) produces, for each i, the same estimate as the centralized Kalman filter (9)-(13).

Proof: Recall that matrices A, Q, C and R are all block diagonal. It then follows from (10) that, if $\Sigma_{k|k}$ is block diagonal, so is $\Sigma_{k+1|k}$. One can easily check from (12) and (13) that the same holds for K_k and $\Sigma_{k|k}$ if $\Sigma_{k|k-1}$ is block diagonal. Since $\Sigma_{1|0} = P$ is block diagonal, it follows that the matrices $\Sigma_{k|k-1}$ and $\Sigma_{k|k}$ remain block diagonal for all k. Now, it is straightforward to verify that (9)-(13) become equivalent to (15)-(19), when $\Sigma_{k|k}$ and $\Sigma_{k|k-1}$ are block diagonal. Hence, the distributed and centralized Kalman filters produce the same estimate and the result follows. ■

VI. Simulations

In this section, we present four numerical experiments to study the convergence of the proposed distributed Kalman filter to its centralized counterpart. In the first experiment, we compare the convergence on networks with different topologies. To this end, we consider a directed communication topology involving ten subsystems with first-order dynamics. The subsystems' initial conditions are drawn from the normal distribution $\mathcal{N}(0, P)$. The initial covariance matrix P is chosen by randomly generating a positive-definite matrix using $P = \mathcal{L}\mathcal{L}^{\top} + \epsilon_0 I_{10}$, with $\epsilon_0 = 0.1$, and the entries of $\mathcal{L} \in \mathbb{R}^{10\times10}$ are drawn from the uniform distribution $\mathcal{U}(0, 1)$. Also, $v_k \sim \mathcal{N}\left(0, 0.1I_{10}\right)$ and $w_k \sim \mathcal{N}\left(0, 0.1I_{10}\right)$. The poles of these ten subsystems are randomly chosen from the uniform distribution $\mathcal{U}(0.4, 0.8)$. We consider two different topologies. The first one is a *path* topology, whose weights, i.e., the scalars $L^{(i,j)}$, are randomly selected from the distribution $\mathcal{U}(0, 1)$. The second one is a *random* topology, whose weights are randomly drawn while guaranteeing the stability of the overall system. We refer to these two topologies as Case A and Case B, respectively. We examine

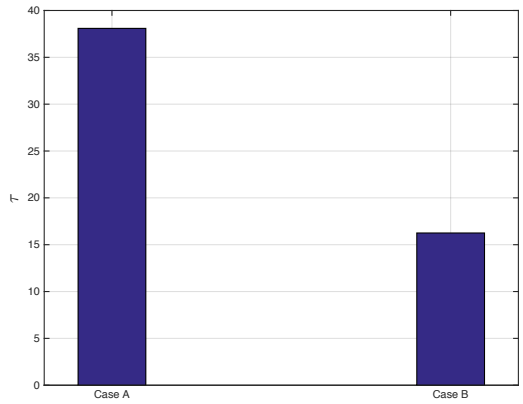

Fig. 1. Settling time vs. different chioces of topologies.

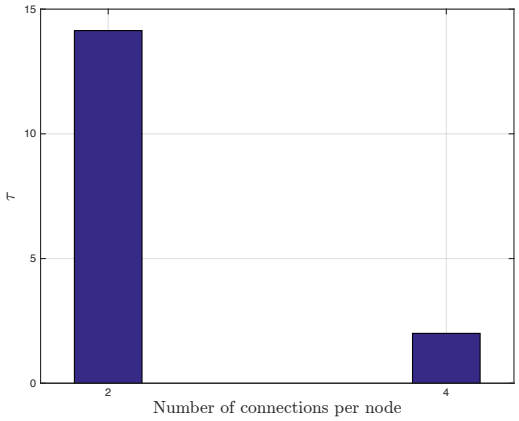

Fig. 2. Settling time vs. number of connections per node.

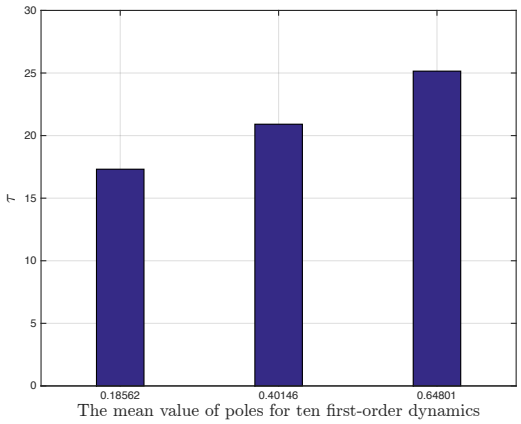

Fig. 3. Settling time vs. mean value of subsystems' poles, in a network with random topology.

the convergence rate of the proposed filtering algorithm for these two cases. To measure this rate, we use the settling time, which we define by

$$\tau = \min_{\tilde{\tau}} \left\{ \max_{k:\tilde{\tau}\leq k\leq T} ||\tilde{x}_{x_{k|k-1}}||_2 < 0.1||\tilde{x}_{0|-1}||_2 \right\}, \quad (31)$$

where T is the running time. For each one of these two topologies, we plot the average value of τ obtained over 200 realizations of the process noise $w_k^{(i)}$, measurement noise $v_k^{(i)}$, initial condition and the random selection of the topology weights. The results are shown in Figure 1. We see that in Case B the convergence is faster. The reason for this is that the random topology generated in Case B has a much larger number of edges compared to the path topology in Case A. Indeed, the matrix L associated with Case B is a very dense one, i.e., it contains very few zero entries.

In the second experiment we study the effect of number of connections per node on the convergence rate. To this end, we consider a network of five nodes with first-order dynamics, poles at 0.15 and topology weights $L^{(i,j)} = 0.1$. We perform 200 Monte Carlo simulations for two classes of topologies, namely, with two and four connections per node. One can observe that the former results in a *cycle* graph and the latter delivers a *complete* graph. The simulation results are depicted in Figure 2. Again, we see how the settling time decreases with the number of connections.

In the third experiment we study the effect of subsystems' dynamics on the settling time of the distributed Kalman filter. We consider a network with ten nodes with first-order dynamics, interconnected with a random topology. We consider three cases, in which the subsystems' poles are drawn from the uniform distributions $\mathcal{U}(0.1, 0.3), \mathcal{U}(0.2, 0.6)$ and $\mathcal{U}(0.6, 0.8)$, respectively. Figure 3 shows the dependence of τ, obtained by averaging 200 Monte Carlo runs, on the mean value of the subsystems' poles. We see how the settling time increases with this value.

In our final experiment, we compare the settling time for networks with different sizes. We consider three cases, having 5, 10 and 20 nodes, respectively. For each case, we consider identical nodes with first-order dynamics, having poles at 0.9, and connected using a loop topology whose nonzero values are $L^{(i,j)} = 0.1$. Figure 4 depicts the dependence of the settling time τ, again obtained by averaging 200 Monte Carlo runs, on the number of nodes. We see how the increase of this time with the network size is only marginal.

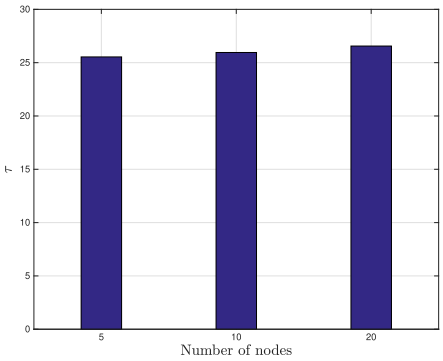

Fig. 4. Settling time vs. number of nodes, in a network with path topology.

VII. Conclusion

We examined the problem of distributed estimation for a network of dynamical subsystems. In particular, we consid-

ered a novel version of the Kalman filter that only exploits local measurements. We then studied the performance of this proposed filter with respect to a traditional Kalman filter that has access to all measurments within the network. In this paper, we illustrated that the covariance matrix associated with the initial value of the state vector plays an important role on the outcomes of the distributed Kalman filter. We showed that if this matrix is block diagonal, the proposed distributed scheme is optimal. Moreover, if that condition is dropped, the estimation error covariances, and the associated estimates, obtained through these two approaches approximate each other exponentially fast. We also established proper bounds on error between estimates and its covariance matrix.

REFERENCES

[1] M. Subbotin and R. Smith, "Design of distributed decentralized estimators for formations with fixed and stochastic communication topologies," *Automatica*, vol. 45, no. 11, pp. 2491 – 2501, 2009.

[2] Z. Lin, L. Wang, Z. Han, and M.Fu, "Distributed formation control of multi-agent systems using complex Laplacian," *IEEE Transactions Automatic Control*, vol. 59, no. 7, pp. 1765–1777, 2014.

[3] W. Zhang, M. S. Branicky, and S. M. Phillips, "Stability of networked control systems," *IEEE Control Systems*, vol. 21, no. 1, pp. 84–99, 2001.

[4] A. Teixeira, I. Shames, H. Sandberg, and K. H. Johansson, "A secure control framework for resource-limited adversaries," *Automatica*, vol. 51, pp. 135–148, 2015.

[5] M. Zamani, U. Helmke, and B. D. O. Anderson, "Zeros of networked systems with time-invariant interconnections,," *Automatica*, pp. 97–105, 2015.

[6] J. A. Marshall, M. E. Broucke, and B. A. Francis, "Formations of vehicles in cyclic pursuit," *IEEE Transactions on Automatic Control*, vol. 49, no. 11, pp. 1963–1974, 2004.

[7] M. A. Bruckstein, G. Sapiro, and D. Shaked, "Evolutions of planar polygons." *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 9, no. 6, pp. 991–1014, 1995.

[8] R. W. Brockett and J. L. Willems, "Discretized partial differential equations: Examples of control systems defined on modules," *Automatica*, vol. 10, no. 5, pp. 507–515, 1974.

[9] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1520–1533, 2004.

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

[11] Z. Lin, L. Wang, Z. Han, and M. Fu, "A graph laplacian approach to coordinate-free formation stabilization for directed networks," *IEEE Transactions Automatic Control*, vol. 61, no. 5, pp. 1269–1280, 2016.

[12] R. Zheng, Z. Lin, M. Fu, and D. Sun, "Distributed control for uniform circumnavigation of ring-coupled unicycles," *Automatica*, vol. 53, pp. 23–29, 2015.

[13] Z. Lin, L. Wang, Z. Chen, M. Fu, and Z. Han, "Necessary and sufficient graphical conditions for affine formation control," *IEEE Transactions Automatic Control*, vol. 61, no. 10, pp. 2877–2891, 2016.

[14] C. G. Lopes and A. H. Ali, "Diffusion least-mean squares over adaptive networks: Formulation and performance analysis," *IEEE T. Signal Proces*, vol. 56, no. 7, pp. 3122–3136, 2008.

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

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

[17] A. Gómez-Expósito, A. D. la Villa Jaén, C. Gómez-Quiles, P. P. Rousseaux, and T. V. Cutsem, "A taxonomy of multi-area state estimation methods," *Electric Power Systems Research*, vol. 81, no. 4, pp. 1060–1069, 2011.

[18] D. Marelli and M. Fu, "Distributed weighted linear least squares estimation with fast convergence in large-scale systems," *Automatica*, vol. 51, pp. 27–39, 2015.

[19] R. Olfati-Saber, "Distributed kalman filter with embedded consensus filters," in *IEEE Conference on Decision and Control*, 2005, pp. 8179–8184.

[20] V. Ugrinovskii, "Distributed robust filtering with consensus of estimates," *Automatica*, vol. 47, no. 1, pp. 1 – 13, 2011.

[21] ——, "Distributed robust estimation over randomly switching networks using consensus," *Automatica*, vol. 49, no. 1, pp. 160 – 168, 2013.

[22] M. Zamani and V. Ugrinovskii, "Minimum-energy distributed filtering," in *IEEE Conference on Decision and Control*, 2014, pp. 3370–3375.

[23] 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, 2008.

[24] R. Olfati-Saber, "Kalman-consensus filter : Optimality, stability, and performance," in *IEEE Conference on Decision and Control*, 2009, pp. 7036–7042.

[25] A. Ribeiro, I. Schizas, S. Roumeliotis, and G. Giannakis, "Kalman filtering in wireless sensor networks," *IEEE Control Systems*, vol. 30, no. 2, pp. 66–86, 2010.

[26] M. Zamani, B. Ninness, and J. C. Agüro, "On identification of networked systems with time-invariant topology," in *IFAC Symposium on System Identification*, 2015, pp. 1184–1189.

[27] B. M. Sanandaji, T. L. Vincent, and M. B. Wakin, "A review of sufficient conditions for structure identification in interconnected systems," in *IFAC Symposium on System Identification*, vol. 45, no. 16, 2012, pp. 1623–1628.

[28] ——, "Exact topology identification of large-scale interconnected dynamical systems from compressive observations," in *American Control Conference*. IEEE, 2011, pp. 649–656.

[29] A. Dankers, P. M. J. V. den Hof, X. Bombois, and P. S. C. Heuberger, "Identification of dynamic models in complex networks with prediction error methods: Predictor input selection," *IEEE Transactions on Automatic Control*, vol. 61, no. 4, pp. 937–952, 2016.

[30] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, NJ: Prentice-Hall, 1979.

[31] P. Bougerol, "Kalman filtering with random coefficients and contractions," *SIAM Journal on Control and Optimization*, vol. 31, no. 4, pp. 942–959, 1993.

[32] T. Sui, D. Marelli, and M. Fu, "Accuracy analysis for distributed weighted least-squares estimation in finite steps and loopy networks," submitted to Automatica. [Online]. Available: http://www.cifasis-conicet.gov.ar/marelli/DWLS-accuracy-full.pdf

[33] D. Marelli, M. Zamani, and M. Fu, "Distributed kalman filter in a network of linear dynamical systems," techical report. [Online]. Available: https://arxiv.org/abs/1711.07625