EUSIPCO 2013 1569744769

DISTRIBUTED KALMAN FILTERING

Subhro Das and José M.F. Moura

Carnegie Mellon University
Department of Electrical and Computer Engineering, Pittsburgh, PA 15213 USA

ABSTRACT

This paper proves the asymptotic convergence of the mean-squared error (MSE) of a distributed Kalman filter that we have previously proposed. This result shows that our distributed Kalman filter can track with bounded MSE any arbitrary linear dynamics.

***Index Terms*—** Distributed Kalman filter, asymptotic convergence, mean-squared error

1. INTRODUCTION

Distributed problems have been extensively studied in the last few years, e.g., [1], [2], [3], [4], to perform distributed averaging, consensus, detection or parameter estimation. Distributed estimation of linear time processes has also been the focus of intense research. In [5] and [6], a distributed version of Kalman filter is proposed where all agents have the same copy of the linear process and run the same local copy of a Kalman filter after performing consensus on the innovations. In [7], the process itself is distributed so that the distributed Kalman filter proposed in this reference is different in different agents, avoiding the problem of agents of limited resources running locally high dimensional Kalman filters. The filter in [7] still performs cycles of consensus in between measurements, see the reference for details.

In [8], a new distributed Kalman filter structure is proposed where sensors run copies of local filters, exchanging their states according to a gossip protocol. This reference shows that under an ergodic Markov schedule (of visiting the agents) the gossip distributed Kalman filter exhibits asymptotically optimal behavior (unbiasedness, consistency, efficiency). In [9] and [10], a distributed Kalman filter is proposed and the concept of Network Tracking Capacity (NTC) is introduced. These references show that the proposed distributed linear estimator can track unstable linear dynamics as long as the degree of instability is bounded by the NTC.

In [11], we have proposed a distributed Kalman filter of the *consensus + innovations* type that extends [4] and [12] to linear dynamic processes. In [11], we proved this linear distributed Kalman estimator is asymptotically unbiased. In this paper we show that it can track with bounded mean-squared error (MSE) any arbitrary unstable dynamics, as long as the ratio of the algebraic connectivity and the largest eigenvalue of the graph Laplacian of the network topology is lower bounded. The bound depends only on the network topology whereas in prior work the NTC depends both on the network topology and the observation model.

The rest of the paper is organized as follows: Section 2 introduces the notations and provides a preliminary background. Section 3 includes the distributed Kalman filtering algorithm. In Section 4, we discuss the different design parameters of the estimator. In Section 5 the convergence of the error process is proved and analyzed its performance. Section 6 concludes the paper.

This work has been supported by AFOSR grant FA95501010291 and by NSF grants CCF1011903 and CCF1018509.

2. PRELIMINARIES

We define $\theta(i) \in \mathbb{R}^M$ as the state vector of a linear dynamical system with discrete time index i:

$$\theta(i+1) = \theta(i) + W_{\text{ph}}\theta(i) + v(i) \tag{1}$$

where, the system matrix is $W_{\text{ph}} \in \mathbb{R}^{M\times M}$ and the input noise in the system $v(i) \sim \mathcal{N}(0, V)$. The initial condition of the state vector is normally distributed, $\theta(0) \sim \mathcal{N}(\bar{\theta}_0, \Sigma_0)$. In most distributed applications, W_{ph} is highly sparse.

Each of the N agents in the network observes only a portion of the state $\theta(i)$. Agent n makes the measurement:

$$z_n(i) = H_n\theta(i) + r_n(i) \tag{2}$$

where: $z_n(i) \in \mathbb{R}^{M_n}$ has dimension $M_n \ll M$. The observation matrix is $H_n \in \mathbb{R}^{M\times M_n}$ and the observation noise is $r_n(i) \sim \mathcal{N}(0, R_n)$. For ease of the presentation, we introduce a vector notation. Let:

$$z(i) \quad = \quad H\theta(i) + r(i) \tag{3}$$

where,

$$z(i) = \begin{bmatrix} z_1(i) \\ \vdots \\ z_N(i) \end{bmatrix},\ H = \begin{bmatrix} H_1 \\ \vdots \\ H_N \end{bmatrix},\ r(i) = \begin{bmatrix} r_1(i) \\ \vdots \\ r_N(i) \end{bmatrix} \tag{4}$$

The observation noise $r(i) \sim \mathcal{N}(0, R)$, where the noise covariance is $R = \text{blockdiag}[R_1, \ldots, R_N]$.

The system noise $\{v(i)\}$, the observation noise $\{r(i)\}$, and the initial state $\{\theta(0)\}$ are assumed to be a Gaussian system of uncorrelated random vectors. The noise sequences $\{v(i)\}$ and $\{r(i)\}$ are statistically time-independent.

2.1. System Observability

Denote the state update matrix in (1) by $A = I_M + W_{\text{ph}}$. Given the dynamical system in (1) and the observation model in (3), a centralized Kalman filter could estimate the system state with bounded error covariance if the (A, H) is observable. Since the pair (A, H) is observable, the Kalman gain $\mathcal{K}$ stabilizes the closed loop matrix and the steady state Kalman filter is asymptotically stable, i.e., the closed loop filter matrix $(A - A\mathcal{K}H)$ has poles inside the unit circle, irrespective of A being or not asymptotically stable. In this paper, we assume the system to be globally observable. Then the observability Grammian

$$\mathcal{O} = [H^T \; A^T H^T \ldots \; (A^{M-1})^T H^T]^T \tag{5}$$

is full-rank and the matrix $\mathcal{O}^T \mathcal{O}$ is invertible. We note that global observability does not imply local observability, i.e., none of the agents is required to be observable.

2.2. Network Topology

The communication among the agents is modeled with the undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. The Laplacian matrix of $\mathcal{G}$ is denoted by L. We assume the network to be connected. Then, the eigenvalues of the positive semi-definite matrix L are $0 = \lambda_1(L) < \lambda_2(L) \leq \ldots \leq \lambda_N(L)$, with the algebraic connectivity $\lambda_2(L)$ to be strictly positive. Refer to [13] for more details on graph properties.

The stability of the dynamical system (1) is characterized by the induced 2-norm of the system matrix $a = \|A\| = \|I_M + W_{\text{ph}}\|$. We will consider (1) to be unstable, i.e., $a > 1$. Define the averaging rate factor γ as the ratio of the algebraic connectivity and the largest eigenvalue of L. In this paper, we make the following assumption on the network topology:

$$\gamma = \frac{\lambda_2(L)}{\lambda_N(L)} \geq \frac{a-1}{a+1} \tag{6}$$

The lower bound (6) on γ makes sure that the agents in the network can fuse their data among themselves at a rate faster than the dynamics of the system. The distributed algorithm discussed in this paper can track any arbitrary linear dynamics if the network topology is designed in such a way that it satisfies assumption (6). This is our version of the NTC introduced in [9, 10].

3. THE DISTRIBUTED ALGORITHM

In this section, we present the distributed algorithm that we proposed in [11] for estimation of any arbitrary linear dynamical system (1). Let $x_n(i)$ be the estimate of the state $\theta(i)$ by agent n at time i given the observations up to time $i-1$. Each agent implements the following distributed Kalman filter:

3.1. Consensus on the local estimates

Each agent combines its estimate $x_n(i)$ with its neighbors' estimates via consensus:

$$\overline{x}_n(i) = x_n(i) - \beta \sum_{l \in \Omega_n} (x_n(i) - x_l(i)) \tag{7}$$

where, β is the consensus weight and Ω_n is the neighborhood of agent n. The initial condition is $x_n(0) = \bar{\theta}_0, \forall n$. In compact vector notation, equation (7) can be written as:

$$\overline{x}(i) = x(i) - \beta(L \otimes I_M)x(i), \quad x(i) = \begin{bmatrix} x_1(i) \\ \vdots \\ x_N(i) \end{bmatrix} \tag{8}$$

The symbol $\otimes$ denotes the Kronecker matrix product. The initial condition becomes $x(0) = 1_N \otimes \bar{\theta}_0$.

3.2. Pseudo-innovations

We define at each agent the pseudo-innovations as:

$$\nu_n(i) = H_n^T \Big(z_n(i) - H_n \overline{x}_n(i) \Big) \tag{9}$$

Each agent estimates the global average $\widehat{\nu}_n(i)$ of the pseudo-innovations. It involves two steps. A consensus step:

$$\overline{\widehat{\nu}}_n(i) = \sum_{l \in \mathcal{N}_n} w_{nl} \widehat{\nu}_l(i), \tag{10}$$

then followed by the update:

$$\widehat{\nu}_n(i+1) = \overline{\widehat{\nu}}_n(i) + (\nu_n(i+1) - C_n \overline{\widehat{\nu}}_n(i)) \tag{11}$$

where, $W = \{w_{nl}\}$ is a fusion weight matrix with the same sparsity as the graph Laplacian L and $\mathcal{N}_n$ is the closed neighborhood of agent n, which includes n also. The initial condition is $\widehat{\nu}_n(0) = \nu_n(0)$. The positive semi-definite matrix C_n has sparsity such that when multiplied with $\overline{\widehat{\nu}}_n(i-1)$ results in a vector with the same sparsity as $\nu_n(i)$.

Equations (9)-(11), in compact notation are:

$$\begin{aligned} \nu(i) &= D_H^T z(i) - \overline{D}_H \overline{x}(i) && (12) \\ \overline{\widehat{\nu}}(i) &= (W \otimes I_M) \widehat{\nu}(i) && (13) \\ \widehat{\nu}(i+1) &= \overline{\widehat{\nu}}(i) + (\nu(i+1) - C\overline{\widehat{\nu}}(i)) && (14) \end{aligned}$$

where, $\nu(i) = [\nu_1^T(i) \;\; \ldots \;\; \nu_N^T(i)]^T$ and

$$\begin{aligned} D_H &= \text{blockdiag}\{H_1, \ldots, H_N\} \\ \overline{D}_H &= \text{blockdiag}\{H_1^T H_1, \ldots, H_N H_N^T\} \\ C &= \text{blockdiag}\{C_1, \ldots, C_N\} \end{aligned}$$

3.3. Predictor

Next each agent combines the state fusion (7) with the pseudo-innovations (11) to obtain its estimate of the current state vector. Each agent then predicts the state vector at time $i+1$ on the basis of the observations up to time i as:

$$x_n(i+1) = (I_M + W_{\text{ph}})(\overline{x}_n(i) + K_n(i)\widehat{\nu}_n(i)) \tag{15}$$

where $K_n(i)$ is the local estimator gain at agent n. In compact notation, equation(15) can be written as:

$$x(i+1) = (I_N \otimes (I_M + W_{\text{ph}}))(\overline{x}(i) + K(i)\widehat{\nu}(i)) \tag{16}$$

where, $K(i) = \text{blockdiag}\{K_1(i), ..., K_N(i)\}$.

In [11], we showed that the distributed Kalman estimator (15) is asymptotically unbiased. Next, we analyze the error process of the distributed estimator (16) and how to design the estimator parameters to obtain bounded MSE.

4. PARAMETER DESIGN

In this section, we discuss the design of three key parameters of the distributed Kalman estimator: the network topology L, the time-varying estimator gain $K(i)$, and the matrix C. The fusion matrix is dependent on the network topology L. We will consider the same fusion weight martix for both state fusion and pseudo-innovations fusion, i.e., $W = I_N - \beta L$. We design the network topology such that:

$$\|A\|\|W\| = a\|I_N - \beta L\| < 1 \tag{17}$$

$$\text{where,}\ \ \|I_N - \beta L\| = \max\{1-\beta\lambda_2(L), \beta\lambda_N(L)-1\} \tag{18}$$

Under assumption (6) on the averaging rate factor γ, β exists and condition (17) is satisfied for all $\beta \in \left(\frac{a-1}{a.\lambda_2(L)}, \frac{a+1}{a.\lambda_N(L)}\right)$. The optimal β^* is given by:

$$\begin{aligned} 1 - \beta^*\lambda_2(L) &= \beta^*\lambda_N(L) - 1 \\ \text{i.e., } \beta^* &= \frac{2}{\lambda_2(L) + \lambda_N(L)} \end{aligned} \tag{19}$$

The above analysis, similar to convergence rate analysis for distributed inference on graphs in [14], assumes that the system (1) is unstable ($a > 1$). If the system is stable, then connectedness of the network ($\lambda_2(L) > 0$) is sufficient for the convergence of the distributed algorithm.

The next design parameter of the estimator is the gain matrix $K(i)$. The initial values and the update equations for the gain matrices $K_n(i)$ at each agent n are:

$$K_n(0) = \left(H_n^T H_n\right)^+ \tag{20}$$

$$K_n(i+1) = \sum_{l \in \mathcal{N}_n} \overline{w}_{nl} K_l(i) \tag{21}$$

where F^+ is the Moore-Penrose pseudo-inverse of the matrix F as defined in [15]. The matrix $\overline{W} = \{\overline{w}_{nl}\} = I_N - \overline{\beta}L$ is the fusion weight matrix for the gain matrices. Note that the gain matrix $K_n(i)$ at each agent converges to the average of the initial gain matrices $K_n(0)$ of all the agents.

Define $P_{NM} = \frac{1}{N}(1_N \otimes I_M)(1_N \otimes I_M)^T$ and note that $P_{NM}(\overline{W} \otimes I_N) = P_{NM}$. It can be shown that:

$$P_{NM}K(i+1) = P_{NM}K(i)(\overline{W} \otimes I_N), \tag{22}$$

from which

$$\begin{aligned}(I_2 \otimes P_{NM}) \begin{bmatrix} I & K(i+1) \\ 0 & I \end{bmatrix} &= \begin{bmatrix} P_{NM} & P_{NM}K(i+1) \\ 0 & P_{NM} \end{bmatrix} \\ &= \begin{bmatrix} P_{NM}(\overline{W} \otimes I_N) & P_{NM}K(i)(\overline{W} \otimes I_N) \\ 0 & P_{NM}(\overline{W} \otimes I_N) \end{bmatrix} \\ &= (I_2 \otimes P_{NM}) \begin{bmatrix} I & K(i) \\ 0 & I \end{bmatrix} (I_2 \otimes (\overline{W} \otimes I_N)) \end{aligned}$$

Taking the 2-norm on both sides, then using the sub-multiplicative property of the matrix norm and properties of the Kronecker product, we get

$$\left\|\begin{matrix} I & K(i+1) \\ 0 & I \end{matrix}\right\| \leq \left\|\begin{matrix} I & K(i) \\ 0 & I \end{matrix}\right\| \|\overline{W}\| \leq \|\overline{W}\|^{i+1} \left\|\begin{matrix} I & K(0) \\ 0 & I \end{matrix}\right\| \tag{23}$$

We design $\overline{\beta}$ such that, for given $K(0)$ and observation model $\overline{D}_H$, there exists some $i_0 > 0$ such that for all $i > i_0$

$$\|\overline{W}\|^i \left\|\begin{matrix} I & K(0) \\ 0 & I \end{matrix}\right\| \left\|\begin{matrix} I & 0 \\ -\overline{D}_H & I \end{matrix}\right\| < 1 \tag{24}$$

and so,

$$\left\|\begin{matrix} I & K(i) \\ 0 & I \end{matrix}\right\| \left\|\begin{matrix} I & 0 \\ -\overline{D}_H & I \end{matrix}\right\| < 1. \tag{25}$$

We will use the above inequality in the next section. Note that the gain terms can be precomputed and stored at each agent to reduce the running complexity of the distributed algorithm. The pseudo-innovations gain matrix C is positive semi-definite and it is designed such that

$$\|I - C\| \leq 1. \tag{26}$$

5. PERFORMANCE ANALYSIS

Define the error process $e(i)$:

$$e(i) = x(i) - 1_N \otimes \theta(i) \tag{27}$$

Using (16) and (1), with $A = I_M + W_{\text{ph}}$,

$$\begin{aligned}
e(i+1) &= x(i+1) - 1_N \otimes \theta(i+1) \\
&= (I_N \otimes A)\Big(\overline{x}(i) + K(i)\widehat{\nu}(i)\Big) - 1_N \otimes \Big(A\theta(i) + v(i)\Big) \\
&= (I_N \otimes A)\Big(\big(W \otimes I_M\big)x(i) - 1_N \otimes \theta(i) + K(i)\widehat{\nu}(i)\Big) \\
&\qquad -1_N \otimes v(i) \\
&= (I_N \otimes A)\Big(\big(W \otimes I_M\big)\big(x(i) - 1_N \otimes \theta(i)\big) + K(i)\widehat{\nu}(i)\Big) \\
&\qquad -1_N \otimes v(i) \\
&= (I_N \otimes A)\Big(\big(W \otimes I_M\big)e(i) + K(i)\widehat{\nu}(i)\Big) - 1_N \otimes v(i) \quad (28)
\end{aligned}$$

We write equation (12) of the pseudo-innovations as follows:

$$\begin{aligned}
\nu(i) &= D_H^T z(i) - \overline{D}_H \overline{x}(i) \\
&= D_H^T D_H \big(1_N \otimes \theta(i)\big) - \overline{D}_H \overline{x}(i) + D_H^T r(i) \\
&= -\overline{D}_H \Big(\big(W \otimes I_M\big)x(i) - 1_N \otimes \theta(i)\Big) + D_H^T r(i) \\
&= -\overline{D}_H \big(W \otimes I_M\big)e(i) + D_H^T r(i) \qquad (29)
\end{aligned}$$

From equation (14), the pseudo-innovations estimates are:

$$\begin{aligned}
\widehat{\nu}(i) &= \overline{\widehat{\nu}}(i-1) + \big(\nu(i) - C\overline{\widehat{\nu}}(i-1)\big) \\
&= \big(I_{NM} - C\big)\big(W \otimes I_M\big)\widehat{\nu}(i-1) \\
&\qquad -\overline{D}_H \big(W \otimes I_M\big)e(i) + D_H^T r(i) \qquad (30)
\end{aligned}$$

Equations (30) and (28) imply:

$$\begin{aligned}
e(i+1) &= (I_N \otimes A)\Big(\big(I_{NM} - K(i)\overline{D}_H\big)\big(W \otimes I_M\big)e(i) \\
&\qquad + K(i)\big(I_{NM} - C\big)\big(W \otimes I_M\big)\widehat{\nu}(i-1)\Big) \\
&\qquad + \big(I_N \otimes A\big)K(i)D_H^T r(i) - 1_N \otimes v(i) \qquad (31)
\end{aligned}$$

In compact form, equations (30) and (31) can be written as:

$$\begin{aligned}
&\begin{bmatrix} e(i{+}1) \\ \widehat{\nu}(i) \end{bmatrix} = \\
&\begin{bmatrix} (I \otimes A)\big(I - K(i)\overline{D}_H\big)\big(W \otimes I\big) & (I_N \otimes A)K(i)\big(I - C\big)\big(W \otimes I\big) \\ -\overline{D}_H\big(W \otimes I\big) & \big(I - C\big)\big(W \otimes I\big) \end{bmatrix} \\
&\quad \times \begin{bmatrix} e(i) \\ \widehat{\nu}(i{-}1) \end{bmatrix} + \begin{bmatrix} \big(I_N \otimes A\big)K(i)D_H^T r(i) - 1_N \otimes v(i) \\ D_H^T r(i) \end{bmatrix} \qquad (32)
\end{aligned}$$

i.e.,

$$\tilde{e}(i{+}1) = \tilde{A}(i)\tilde{e}(i) + \eta(i) \qquad (33)$$

where,

$$\begin{aligned}
\tilde{e}(i) &= \begin{bmatrix} e(i) \\ \widehat{\nu}(i{-}1) \end{bmatrix}, & (34) \\
\eta(i) &= \begin{bmatrix} \big(I_N \otimes A\big)K(i)D_H^T r(i) - 1_N \otimes v(i) \\ D_H^T r(i) \end{bmatrix}, & (35)
\end{aligned}$$

$$\begin{aligned}
\tilde{A}(i) &= \begin{bmatrix} (I \otimes A)\big(I - K(i)\overline{D}_H\big)\big(W \otimes I\big) & (I \otimes A)K(i)\big(I - C\big)\big(W \otimes I\big) \\ -\overline{D}_H\big(W \otimes I\big) & \big(I - C\big)\big(W \otimes I\big) \end{bmatrix} \\
&= \begin{bmatrix} I \otimes A & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} I - K(i)\overline{D}_H & K(i) \\ -\overline{D}_H & I_{NM} \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & I - C \end{bmatrix} \\
&\qquad \times \begin{bmatrix} W \otimes I & 0 \\ 0 & W \otimes I \end{bmatrix} \\
&= \begin{bmatrix} I \otimes A & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} I & K(i) \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ -\overline{D}_H & I \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & I - C \end{bmatrix} \\
&\qquad \times \begin{bmatrix} W \otimes I & 0 \\ 0 & W \otimes I \end{bmatrix}
\end{aligned}$$

where, the last step is performed using Schur complement. Taking the 2-norm on both sides, then using the sub-multiplicative property of the matrix norm and properties of the Kronecker product, we get

$$\big\|\tilde{A}(i)\big\| \leq \|A\| \left\| \begin{matrix} I & K(i) \\ 0 & I \end{matrix} \right\| \left\| \begin{matrix} I & 0 \\ -\overline{D}_H & I \end{matrix} \right\| \left\| \begin{matrix} I & 0 \\ 0 & I - C \end{matrix} \right\| \|W\|$$

This and equations (17), (25), and (26) imply,

$$\big\|\tilde{A}(i)\big\| < 1, \quad \forall i \geq i_0. \qquad (36)$$

To analyze the mean-squared boundedness of the estimator (16), we need to analyze the error process (33). Define $\phi(i)$ as the noise covariance matrix of the noise process $\eta(i)$ in the error process $\tilde{e}(i)$. The noise process $\eta(i)$ is a linear combination of the system noise and the observation noise. Note that it satisfies:

$$\begin{aligned}
&\mathbb{E}\left[\eta(i)\right] = 0, \quad \mathbb{E}\left[\eta(i)\eta^T(j)\right] = 0, \quad i \neq j & (37) \\
&\mathbb{E}\left[\eta(i)\eta^T(i)\right] = \phi(i) & (38)
\end{aligned}$$

The time dependency of $\phi(i)$ comes from the time-varying gain matrix $K(i)$. Since the elements of the gain matrix are designed to converge, see (21), there exists an upper bound Φ:

$$\big\|\phi(i)\big\| \quad \leq \quad \Phi, \quad \forall i. \qquad (39)$$

Define the error covariance matrix $S(i+1)$ as:

$$\begin{aligned}
S(i+1) &= \mathbb{E}\left[\tilde{e}(i+1)\tilde{e}^T(i+1)\right] = \tilde{A}(i)S(i)\tilde{A}^T(i) + \phi(i) \\
&= \pi_{0,i}S(0)\pi_{0,i}^T + \sum_{j=0}^{i} \pi_{j,i}\phi(j)\pi_{j,i}^T \qquad (40)
\end{aligned}$$

where, $\qquad \pi_{j,i} = \prod_{k=j}^{i} \tilde{A}(k).$

Since $\big\|\tilde{A}(i)\big\| < 1, \quad \forall i \geq i_0,\ \pi_{0,\infty} = \mathbf{0},$ and so:

$$S(\infty) = \lim_{k\to\infty} s(i+1) = \sum_{j=0}^{\infty} \pi_{j,\infty}\phi(j)\pi_{j,\infty}^T \tag{41}$$

Taking 2-norm on both sides,

$$\begin{aligned} \|S(\infty)\| &\leq \sum_{j=0}^{\infty} \|\pi_{j,\infty}\phi(j)\pi_{j,\infty}^T\| \leq \sum_{j=0}^{\infty} \|\pi_{j,\infty}\|^2 \|\phi(j)\| \\ &\leq \Phi \sum_{j=0}^{\infty} \|\pi_{j,\infty}\|^2 = \Phi \sum_{j=i_0}^{\infty} \|\pi_{j,\infty}\|^2 \\ &\leq \Phi \sum_{j=i_0}^{\infty} \prod_{k=j}^{\infty} \|\tilde{A}(k)\| \leq \Phi \sum_{j=i_0}^{\infty} \prod_{k=j}^{\infty} \|\tilde{A}(i_0)\|^2 \\ &\leq \Phi(1+\rho^2+\rho^4+\rho^6+\ldots) \\ &\leq \frac{\Phi}{1-\rho^2} \end{aligned} \tag{42}$$

where, $\rho = \|\tilde{A}(i_0)\| < 1$. As it can be seen in (42), the 2-norm of the steady state error covariance is upper bounded. The upper bound depends on the 2-norm of the $\tilde{A}(i_0)$ matrix and the 2-norm of the noise covariance matrix of the error process.

6. CONCLUSIONS

In this paper, we have designed a new class of single time-scale distributed estimation algorithms of the Kalman filter type that can track any arbitrary linear dynamical system with bounded mean-squared error. We analyzed the performance of the distributed Kalman estimator. The mean-squared boundedness of our estimation algorithm depends on the lower boundedness of the graph Laplacian of the multi-agent network topology and is independent of the observation model, as long as the system is globally observable. This contrasts with prior work [9] and [10] where the boundedness of the distributed estimator depends on both the network topology and the observation model.

7. REFERENCES

[1] V.D. Blondel, J.M. Hendrickx, A. Olshevsky, and J.N. Tsitsiklis, "Convergence in multiagent coordination, consensus, and flocking," in *44th IEEE Conference on Decision and Control, and European Control Conference, CDC-ECC'05*, 2005, pp. 2996–3000.

[2] S.A. Aldosari and J.M.F. Moura, "Distributed detection in sensor networks: Connectivity graph and small world networks," in *Conference Record of the Thirty-Ninth Asilomar Conference on Signals, Systems and Computers*, 2005, pp. 230–234.

[3] Shreevatsa Rajagopalan and Devavrat Shah, "Distributed averaging in dynamic networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 4, pp. 845–854, 2011.

[4] 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.

[5] Reza Olfati-Saber, "Distributed Kalman filter with embedded consensus filters," in *44th IEEE Conference on Decision and Control and European Control Conference, CDC-ECC'05*, 2005, pp. 8179–8184.

[6] R. Olfati-Saber, "Distributed Kalman filtering for sensor networks," in *46th IEEE Conference on Decision and Control*, 2007, pp. 5492–5498.

[7] 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.

[8] S. Kar and J.M.F. Moura, "Gossip and distributed Kalman filtering: Weak consensus under weak detectability," *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1766–1784, 2011.

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

[10] U.A. Khan and A. Jadbabaie, "Networked estimation under information constraints," *arXiv preprint arXiv:1111.4580*, 2011.

[11] S. Das and J.M.F. Moura, "Distributed state estimation in multi-agent networks," in *38th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2013.

[12] S. Kar and J.M.F. Moura, "Convergence rate analysis of distributed gossip (linear parameter) estimation: Fundamental limits and tradeoffs," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 4, pp. 674–690, 2011.

[13] F.R.K. Chung, *Spectral Graph Theory*, vol. 92, American Mathematical Society, 1997.

[14] S. Kar, S. Aldosari, and J.M.F. Moura, "Topology for distributed inference on graphs," *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2609–2613, 2008.

[15] A. Ben-Israel and T.N.E. Greville, *Generalized inverses: theory and applications*, vol. 15, Springer, 2003.